

# Selecting Algorithms for Black Box Matrices: Checking for Matrix Properties That Can Simplify Computations

Wayne Eberly  
 Department of Computer Science  
 University of Calgary  
 eberly@ucalgary.ca

November 1, 2016

## Abstract

Processes to automate the selection of appropriate algorithms for various matrix computations are described. In particular, processes to check for, and certify, various matrix properties of black-box matrices are presented. These include sparsity patterns and structural properties that allow “superfast” algorithms to be used in place of black box algorithms. Matrix properties that hold generically, and allow the use of matrix preconditioning to be reduced or eliminated, can also be checked for and certified — notably including in the small-field case, where this presently has the greatest impact on the efficiency of the computation.

## 1 Introduction

Krylov-based “black box” algorithms for matrix computations have been used for significant applications in computational number theory. They also form a significant part of the C++ template library LINBOX for high-performance matrix computations. These are notable, in part, because of their versatility: Any matrix representation that allows the input matrix (or, for some algorithms, its transpose) to be multiplied by a vector can be supported.

Considerably more efficient algorithms can be used instead if the input matrix is sparse, with nonzero entries limited to specific locations, or satisfies one of various structural properties. As described, for example, by Golub and van Loan [6], Gaussian Elimination can be used quite efficiently to solve banded systems of linear equations. As described by Pan [7], various classes of matrices (including Toeplitz-like and Hankel-like matrices) have various displacement structures that can be used to reduce system solving for these matrices to polynomial arithmetic. Under these circumstances, assistance in selecting algorithms to be employed might be of help as the community of users of systems like LINBOX grows and non-expert users should be better supported.

Property	Detection/Certification	Verification	Communication
Band Matrix	$\mu k$	$nk + \mu$	$nk$
Low Displacement Rank	$nk^2 + \mu k \log n$	$nk + n \log n + \mu$	$nk + n \log n$
Few Nilpotent Blocks	$n^2 k + \mu n$	$nk + \mu$	$nk$
Many Nilpotent Blocks	$n^2 k + \mu n$	$nk + \mu$	$nk$
Few Invariant Factors	$n^2 \mathcal{M}(n) + n^2 k \log n$	$n \mathcal{M}(n) + nk + \mu$	$nk + n \log n$
Many Invariant Factors	$n^2 \mathcal{M}(n) + n^2 k \log n$ $+ \mu n \log n$	$n \mathcal{M}(n) + nk + \mu$	$nk + n \log n$

Table 1: Summary of Results

Sections 2–4 of this report therefore concern attempts to detect and certify matrix properties to facilitate algorithm selection. The preliminary results given here establish that band matrices and matrices with low displacement rank — including Toeplitz-like and Hankel-like matrices — are easily detected and certified. Furthermore, it is possible to convert matrix representations, in order to allow superfast algorithms to be applied, when such matrices are discovered.

As black-box algorithms have been developed, several matrix properties have been identified that hold generically and that be exploited — generally by eliminating “matrix preconditioning” — to simplify or accelerate computations without sacrificing reliability. In particular, the cost of system solving can be reduced if the input matrix has a small number of nontrivial nilpotent blocks in its Jordan normal form. Simpler algorithms to compute the rank or characteristic polynomial of a matrix can be applied if the input matrix has a small number of nontrivial invariant factors. Sections 5–7 concern the detection and certification of these properties. A technique of Villard [8] is adapted, for the small-field case, to efficiently check for these properties at a cost linear in that needed to apply Wiedemann’s algorithm to compute the minimal polynomial of a matrix. Interactive protocols, of the type recently described by Dumas and Kaltofen [1] are also provided.

Of course, many randomized black box algorithms are Las Vegas, so that one can simply execute algorithms without preconditioning, in hopes that desirable matrix properties are satisfied or that one “gets lucky”. The above results may nevertheless be of interest if one considers exchanges between a service provider and client involving the cost of a service that is to be provided: One would hope here that the cost to the service provider (or “prover”) would not exceed the lower cost to carry out a computation without preconditioning, while the cost to the client (or “verifier”) would be significantly lower than that. Furthermore, a process to certify that preconditioning is necessary would also be of interest. Protocols to certify this are also given.

The expected (and, in a few cases, worst-case) costs established to detect and certify these properties, and for their verification, are linear in the expressions shown in Table 1. In each case, the indicated cost is the number of field operations required to carry out the

indicated operation for a black-box matrix  $A \in \mathbb{F}^{n \times n}$ . Here,  $\mu \leq n(2n + 1)$  is the number of operations required to multiply either  $A$  or  $A^T$  by a given vector  $v \in \mathbb{F}^{n \times 1}$ . It is also assumed that  $\mu \geq n$ .  $\mathcal{M}(n)$  is the number of operations in  $\mathbb{F}$  required for arithmetic in an extension  $\mathbb{E}$  of  $\mathbb{F}$  with logarithmic degree, so that  $\mathcal{M}(n) \in O(\log_2 n \log_2 \log_2 n \log_2 \log_2 \log_2 n)$ . The “communication” reports the number of elements of the ground field  $\mathbb{F}$  (or, in some cases, bits) that must be communicated between a prover and a verifier — excluding the initial cost to communicate a black box matrix  $A \in \mathbb{F}^{n \times n}$ , parameter  $k$  and error tolerance  $\epsilon$ . In typical applications one might expect  $k$  to be significantly smaller than  $n$  — indeed, polylogarithmic. The cost to check for band structure or low displacement rank, and return the information needed for superfast algorithms to be applied when they can, is significantly dominated by the cost to use a black box algorithm to complete a computation instead, in this case.

A conference report [4], including the material found in this report but that omits the proof of Theorem 5.2, is also available.

## 2 Band Structure

Let  $A \in \mathbb{F}^{n \times n}$  and let  $k$  be a positive constant. Let us say that  $A$  is a **band matrix** with **band width**  $k$  if the entry  $a_{i,j}$  of  $A$  in row  $i$  and column  $j$  is equal to 0 whenever  $1 \leq i, j \leq n$  and  $|i - j| > k$ . Golub and van Loan [6] describe efficient algorithms, based on Gaussian Elimination, for computations on such matrices.

As shown below the detection and certification of a black-box matrix that is a band matrix, and conversion to a representation allowing other algorithms to be used, is surprisingly easy. Indeed, this is included, in part, to provide a very simple first example.

The objective of this section is to prove the following.

**Theorem 2.1.** *It is possible to check whether a matrix  $A \in \mathbb{F}^{n \times n}$  is a band matrix, with band width  $k$ , by selecting  $n$  values uniformly and independently from a finite subset  $S$  of  $\mathbb{F}$  and by performing  $\Theta(\mu k)$  arithmetic operations over  $\mathbb{F}$ . If  $A$  is, indeed, a band matrix, then this is confirmed with certainty. Otherwise the probability that  $A$  is mistaken for a band matrix is at most  $1/|S|$ .*

*A certificate with size  $\Theta(nk)$  — which also allows algorithms for band matrix computations to be applied to  $A$  — can be supplied when  $A$  is a band matrix. This certificate can be verified by selecting  $n$  entries uniformly and independently from a set  $S$ , as above, and using  $\Theta(nk + \mu)$  arithmetic operations over  $\mathbb{F}$ . Once again if  $A$  is, indeed, a band matrix and the certificate is correct, then it is accepted with certainty. If the certificate is incorrect then it is accepted with probability at most  $1/|S|$ .*

## 2.1 Detection and Certification

Let  $K = 2k + 1$  and, for  $1 \leq i \leq K$ , let  $\alpha_{K,i} \in \mathbb{F}^{n \times 1}$  such that, for  $1 \leq j \leq n$ , the  $j^{\text{th}}$  entry of  $\alpha_{K,i}$  is equal to one if  $j \equiv i \pmod K$  and is zero otherwise. If  $A$  has band width  $k$  then no two (or more) of columns  $i, i + K, i + 2K, \dots$  of  $A$  have nonzero entries in the same row. The nonzero entries of these columns can therefore simply be read off as entries of the vector  $A \cdot \alpha_{K,i}$  — and all of the nonzero entries of  $A$  can be read off the nonzero entries of each of the vectors  $A \cdot \alpha_{K,1}, A \cdot \alpha_{K,2}, \dots, A \cdot \alpha_{K,K}$ . In particular, if  $1 \leq j \leq n$ , then the  $j^{\text{th}}$  column of  $A$  can only have nonzero entries in rows  $s, s + 1, s + 2, \dots, t$ , where  $s = \max(j - k, 0)$  and  $t = \min(j + k, n)$  — and these entries are the entries of the vector  $A\alpha_{K,j \bmod K}$  in positions  $s, s + 1, s + 2, \dots, t$ .

Consequently,  $K = 2k + 1$  multiplications of  $A$  by vectors suffice for the prover to produce a representation of  $A$  as a band matrix with band width  $k$  if it indeed has this structure.

Of course, this should not be delivered to a verifier without being checked. It is possible that there is no band matrix  $\hat{A} \in \mathbb{F}^{n \times n}$ , with band width  $k$ , such that  $A\alpha_{K,i} = \hat{A}\alpha_{K,i}$  for  $1 \leq i \leq K$  — for it may be necessary for a matrix to have off-band entries in some of columns  $k + 1, k + 2, \dots, K$  or  $n - K, n - K + 1, \dots, n - k - 1$  in order for it to satisfy these equations. In particular (when  $K \leq n$ ), this is the case if the top entry of  $A\alpha_i$  is nonzero for any integer  $i$  such that  $k + 1 \leq i \leq K$ , or if the bottom entry of  $A\alpha_i$  is nonzero for any integer  $i$  such that  $1 \leq i \leq K$  and  $i \in \{n - K + 1 \bmod K, n - K + 2 \bmod K, \dots, n - k - 1 \bmod K\}$ .

Otherwise the band matrix  $\hat{A}$  that satisfies these conditions is unique and it suffices to check that  $A = \hat{A}$ . An application of the test of Frievalds [5] suffices to check this: A vector  $x \in \mathbb{F}^{n \times 1}$ , whose entries are chosen uniformly and independently from a finite subset  $S$  of  $\mathbb{F}$ , should be selected by the prover, and it should be checked whether  $Ax = \hat{A}x$ . If this is not the case then  $A \neq \hat{A}$  and, once again, one should stop.

On the other hand, it is easily checked that if  $A \neq \hat{A}$ , and  $x$  is chosen as above, then the probability that  $Ax = \hat{A}x$  is at most  $1/|S|$  — so that, provided that  $S$  is sufficiently large, the prover should deliver a certificate so that a verification stage can proceed.

## 2.2 Verification

The certificate provided to the verifier, at this point, should simply be a representation of  $A$  as a band matrix — presumably provided as an  $n \times (2k + 1)$  array reporting the entries within the bands of  $A$ .

This can be verified using an independent repetition of the Frievalds test described above.

Since the product of a band matrix  $A \in \mathbb{F}^{n \times n}$  (with band width  $k$ ) and a vector  $x \in \mathbb{F}^{n \times 1}$  can be computed using  $\Theta(nk)$  field operations and zero tests, Theorem 2.1 is now immediate — assuming, again, that  $\mu \geq n$ .

### 3 Low Matrix Rank

The following is less general than the protocol of Dumas and Kaltofen [1] to certify matrix rank and, therefore, inferior in at least one significant respect. However, it can be used in the special case needed here: One is certifying that the rank of  $A \in \mathbb{F}^{n \times n}$  is at most  $k$ , when  $k$  is significantly smaller than  $n$ . It also includes the construction of an alternative representation of  $A$  as needed to support the claims in Section 4.

Suppose now that  $A \in \mathbb{F}^{n \times n}$  has positive rank  $r \leq k$ . Then there exist permutation matrices  $P, Q \in \mathbb{F}^{n \times n}$ , matrices  $L \in \mathbb{F}^{(n-r) \times r}$  and  $R \in \mathbb{F}^{r \times (n-r)}$ , and a nonsingular matrix  $C \in \mathbb{F}^{r \times r}$ , such that

$$A = P \times \begin{bmatrix} I_r \\ L \end{bmatrix} \times C \times \begin{bmatrix} I_r & R \end{bmatrix} \times Q. \quad (3.1)$$

The objective of this section is to establish the following.

**Lemma 3.1.** *It is possible to check whether a matrix  $A \in \mathbb{F}^{n \times n}$  has rank at most  $k$ , by selecting  $\Theta(nk)$  values uniformly and independently from a finite subset  $S$  of  $\mathbb{F}$  and performing  $(nk^2 + \mu k \log n)$  arithmetic operations in  $\mathbb{F}$  and  $\Theta(n \log_2 n)$  operations on bits. This process fails with probability at most  $(\min(r, k) + 1)/|S|$ , where  $r$  is the rank of  $A$ , and only by returning an estimate of the rank of  $A$  that is too low.*

*If  $A$  has rank at most  $k$  then a decomposition of  $A$ , as shown at line (3.1), can be computed at the above cost and returned as a certificate. This certificate can be verified by choosing  $n$  values uniformly and independently from a finite subset  $S$  of  $\mathbb{F}$  and performing  $\Theta(nk + \mu)$  arithmetic operations in  $\mathbb{F}$  and  $\Theta(n \log_2 n)$  operations on bits. If the certificate is correct then it is accepted with certainty. Otherwise it is accepted with probability at most  $1/|S|$ .*

#### 3.1 Detection and Certification

Since the rank of  $A$  cannot exceed that of  $C$ , no decomposition as shown at line (3.1) can exist unless the rank of  $A$  is at most  $k$ . A prover can check for this condition by attempting to construct the matrices included in this decomposition, along with  $C^{-1}$ .

Let  $0 \leq \ell \leq k$  and suppose indices  $i_1, i_2, \dots, i_\ell$  of rows and  $j_1, j_2, \dots, j_\ell$  of columns of an  $\ell \times \ell$  nonsingular submatrix  $C_\ell$  of  $A$  have been computed, along with the matrix  $C_\ell^{-1}$ .

If  $\ell = 0$  then the prover should begin by generating  $n$  values uniformly and independently from a finite subset  $S$  of  $\mathbb{F}$  and using these as the entries of a vector  $x \in \mathbb{F}^{n \times 1}$ . If  $A$  is nonzero then  $Ax \neq 0$  with probability at least  $1 - 1/|S|$  — so that (if  $|S|$  is sufficiently large) the prover may conclude that the rank of  $A$  is zero, if  $Ax = 0$ , and proceed to delivery of a certificate.

Otherwise  $x$  should be used to locate a nonzero column of  $A$ . Suppose that  $x$  has  $h$  nonzero entries. Set  $x_1, x_2 \in \mathbb{F}^{n \times 1}$  such that  $x_1$  has  $\lceil h/2 \rceil$  nonzero entries,  $x_2$  has  $\lfloor h/2 \rfloor$  nonzero entries, and  $x = x_1 + x_2$ . One should check whether  $Ax_1 \neq 0$  — replacing  $x$  with

$x_1$  if this is the case, and replacing  $x$  with  $x_2$  otherwise, since  $Ax_2 = Ax \neq 0$  in this second case. Iterating this process at most  $\lceil \log_2 n \rceil$  times, a vector  $x \in \mathbb{F}^{n \times 1}$  such that  $Ax \neq 0$ , and  $x$  has a single nonzero entry in some position  $j_1$ , is obtained — establishing that the  $j_1^{\text{th}}$  column of  $A$  is nonzero. This column has now been computed, as  $Ax$ , and  $i_1$  can be chosen to be any integer such that  $1 \leq i_1 \leq n$  and the entry  $\alpha$  of  $A$  in row  $i_1$  and column  $j_1$  is nonzero. Now

$$C_1 = [\alpha] \quad \text{and} \quad C_1^{-1} = [\alpha^{-1}].$$

If  $\ell > 0$  then the prover should begin, once again, by forming the vector  $x$  as described above. The prover should continue by computing the matrix-vector product  $v = Ax$ , and setting  $y \in \mathbb{F}^{\ell \times 1}$  to be the vector such that, for  $1 \leq h \leq \ell$ , the entry of  $y$  in position  $h$  is the entry of  $v$  in position  $i_h$ .

The vector  $z = C_\ell^{-1}y \in \mathbb{F}^{\ell \times 1}$  should next be computed. Let  $w \in \mathbb{F}^{n \times 1}$  be the vector such that, for  $1 \leq h \leq \ell$ , the entry of  $w$  in position  $j_h$  is the entry of  $z$  in position  $h$ , and such that all other entries of  $w$  are zero. Finally, set  $u = v - A \cdot w$  — noting that  $v$  is in the space spanned by columns  $j_1, j_2, \dots, j_\ell$  of  $A$  if and only if  $u = 0$ .

If the rank of  $A$  is equal to  $\ell$  then  $u$  must always be equal to zero;  $u$  is nonzero with probability at least  $1 - 1/|S|$  otherwise. Consequently if  $u = 0$  then the prover should proceed with the completion of a certificate, as described below.

Otherwise, if  $x$  has  $h$  nonzero entries then one should once again set  $x_1, x_2 \in \mathbb{F}^{n \times 1}$  such that  $x_1$  has  $\lceil h/2 \rceil$  entries,  $x_2$  has  $\lfloor h/2 \rfloor$  entries, and  $x = x_1 + x_2$ . The above process should be applied to  $x_1$  (instead of  $x$ ) to check whether  $Ax_1$  is in the space spanned by columns  $j_1, j_2, \dots, j_\ell$  of  $A$  — replacing  $x$  with  $x_1$  if this is not the case, and replacing  $x$  with  $x_2$  otherwise. Iterating this process at most  $\lceil \log_2 n \rceil$  times one eventually obtains a vector  $x \in \mathbb{F}^{n \times 1}$  such that  $Ax$  is not in the space spanned by columns  $j_1, j_2, \dots, j_\ell$  of  $A$  and  $x$  has a single nonzero entry in some position  $j_{\ell+1}$ . This establishes that the  $j_{\ell+1}^{\text{th}}$  column of  $A$  is not in the space spanned by columns  $j_1, j_2, \dots, j_\ell$  — and that columns  $j_1, j_2, \dots, j_{\ell+1}$  of  $A$  are linearly independent.

One should next compute the vector  $u \in \mathbb{F}^{n \times 1}$ , as described above, corresponding to the final choice of the vector  $x$  — so that  $u \neq 0$ . It suffices to choose  $i_{\ell+1}$  such that  $1 \leq i_{\ell+1} \leq n$  and the  $i_{\ell+1}^{\text{th}}$  entry of  $u$  is nonzero in order to ensure that the submatrix  $C_{\ell+1}$  of  $A$ , including entries in rows  $i_1, i_2, \dots, i_{\ell+1}$  and columns  $j_1, j_2, \dots, j_{\ell+1}$ , is nonsingular.

Note next that

$$C_{\ell+1} = \begin{bmatrix} C_\ell & s \\ t & \alpha \end{bmatrix}$$

for vectors  $s \in \mathbb{F}^{\ell \times 1}$  and  $t \in \mathbb{F}^{1 \times \ell}$ , and for some value  $\alpha \in \mathbb{F}$ . Since  $C_\ell$  is nonsingular,

$$C_{\ell+1} = \begin{bmatrix} I_\ell & 0 \\ tC_\ell^{-1} & 1 \end{bmatrix} \cdot \begin{bmatrix} C_\ell & 0 \\ 0 & \beta \end{bmatrix} \cdot \begin{bmatrix} I_\ell & C_\ell^{-1}s \\ 0 & 1 \end{bmatrix}$$

where  $\beta = \alpha - tC_\ell^{-1}s$ . Now  $\beta \neq 0$ , since  $C_{\ell+1}$  is also nonsingular, and

$$\begin{aligned} C_{\ell+1}^{-1} &= \begin{bmatrix} I_\ell & -C_\ell^{-1}s \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} C_\ell^{-1} & 0 \\ 0 & \beta^{-1} \end{bmatrix} \cdot \begin{bmatrix} I_\ell & 0 \\ -tC_\ell^{-1} & 1 \end{bmatrix} \\ &= \begin{bmatrix} C_\ell^{-1} + (C_\ell^{-1}s) \cdot (\beta^{-1}tC_\ell^{-1}) & -C_\ell^{-1}s\beta^{-1} \\ -\beta^{-1}tC_\ell^{-1} & \beta^{-1} \end{bmatrix}. \end{aligned}$$

Since  $C_\ell^{-1}s \in \mathbb{F}^{\ell \times 1}$  and  $\beta^{-1}tC_\ell^{-1} \in \mathbb{F}^{1 \times \ell}$ , this expression for  $C_{\ell+1}^{-1}$  can be used to compute the entries of  $C_{\ell+1}^{-1}$  using  $\Theta(\ell^2)$  operations in  $\mathbb{F}$ . The value  $\ell$  can now be incremented and the above process repeated.

If this process is iterated until  $\ell = k + 1$ , then the rank of  $A$  is greater than  $k$  and one can stop. Otherwise the rank  $r \leq k$  of  $A$  has been obtained, along with the matrix  $C = C_r \in \mathbb{F}^{r \times r}$  shown at line (3.1), and the indices of the rows and columns of this matrix in  $A$ .

The permutation matrices  $P$  and  $Q$ , shown at line (3.1), can each be concisely represented as an integer vector, with length  $n$ , whose  $i^{\text{th}}$  entry is the index of the nonzero entry in row  $i$  of the permutation matrix. Since the first  $r$  entries of this representation of  $Q$  are the indices  $j_1, j_2, \dots, j_r$ , it is not difficult to compute this representation of  $Q$  using  $O(n)$  operations on integers whose binary representations have length  $O(\log n)$ : The only operations required are the initialization of these vectors, comparisons of integers and assignments of values. If a first array is initially sorted and a second integer array is used to maintain the locations of each of  $1, 2, \dots, n$  in the initial array then one can reorder  $1, 2, \dots, n$  in order to obtain this representation of  $Q$  using  $O(r)$  exchanges of values in this array. The second array, mentioned above, is then a representation of  $Q^T$ . Since the initial entries of a representation of  $P^T$  are the indices  $i_1, i_2, \dots, i_r$ , a representation of  $P^T$  can be computed in the same way using  $O(n)$  operations on integers with length in  $O(\log n)$ . A representation of  $(P^T)^T = P$  is also obtained as a result of this process.

It remains only to notice that if  $A_L \in \mathbb{F}^{n \times r}$  is the matrix including columns  $j_1, j_2, \dots, j_r$  of  $A$ , and  $A_R \in \mathbb{F}^{r \times n}$  is the matrix including rows  $i_1, i_2, \dots, i_r$  of  $A$ , then

$$\begin{bmatrix} I_r \\ L \end{bmatrix} = P^T \cdot A_L \cdot C^{-1} \quad \text{and} \quad \begin{bmatrix} I_r & R \end{bmatrix} = C^{-1} \cdot A_R \cdot Q^T.$$

Since  $C^{-1}$  has already been computed,  $L$  and  $R$  can be computed using  $O(nr^2)$  additional arithmetic operations in  $\mathbb{F}$  and  $\Theta(n \log_2 n)$  operations on bits.

A consideration of the above confirms that  $\Theta(nk^2 + k\mu \log n)$  arithmetic operations in  $\mathbb{F}$  and  $\Theta(n \log_2 n)$  operations on bits have been used, in the worst case, to check whether the rank of  $A$  is at most  $k$ , and to compute the rank and the decomposition at line (3.1) if this is the case. This process can only fail due to unlucky choices of the randomly selected vectors  $x \in \mathbb{F}^{n \times 1}$ , described above. Since each selection fails with probability at most  $1/|S|$  and at most  $\min(r, k) + 1$  such vectors must be selected if  $A$  has rank  $r$ , the total probability of failure is at most  $(r + 1)/|S|$  if  $r \leq k$  and at most  $(k + 1)/|S|$  otherwise.

### 3.2 Verification

Once again, it suffices to apply the Frievalds test to verify that  $A$  is the zero matrix, if the reported rank is zero, or that the decomposition of  $A$ , shown at line (3.1), is correct otherwise. An examination of this decomposition confirms that this test can be carried out at the cost stated in the above lemma.

## 4 Low Displacement Rank

For  $\alpha \in \mathbb{F}$ , the  $n \times n$   **$\alpha$ -circulant matrix**  $Z_\alpha$  is the matrix

$$Z_\alpha = \begin{bmatrix} 0 & & & & \alpha \\ 1 & 0 & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 0 & \\ & & & 1 & 0 \end{bmatrix} \in \mathbb{F}^{n \times n}$$

whose entry in row  $i + 1$  and column  $i$  is 1 for  $1 \leq i \leq n - 1$ , whose entry in row 1 and column  $n$  is  $\alpha$ , and all of whose other entries are zero. Consider the following linear operators on matrices in  $\mathbb{F}^{n \times n}$ :

- $\varphi_T(A) = Z_1 \cdot A - A \cdot Z_0$ .
- $\varphi_H(A) = Z_1 \cdot A - A \cdot Z_0^T$ .
- $\varphi_{TH}(A) = (Z_0 + Z_0^T) \cdot A - A \cdot (Z_0 + Z_0^T)$ .

A matrix  $A \in \mathbb{F}^{n \times n}$  is **Toeplitz-like** (respectively, **Hankel-like**, and **Toeplitz+Hankel-like**) if the rank of the matrix  $\varphi_T(A)$  (respectively,  $\varphi_H(A)$ , and  $\varphi_{TH}(A)$ ) is small relative to  $n$ . The matrix  $\varphi_T(A)$  (respectively,  $\varphi_H(A)$  or  $\varphi_{TH}(A)$ ) is called the **operator matrix** and rank of this matrix is said to be the **displacement rank** of  $A$ . As described, for example, by Pan [7], a variety of matrix computations have “superfast algorithms” if the displacement rank of a matrix is low. Indeed, if the displacement rank is polylogarithmic in  $n$  then the worst-case running times of these algorithms are generally within a polylogarithmic factor of linear in  $n$ .

A black box for multiplication of  $\varphi_T(A)$  (respectively,  $\varphi_H(A)$  or  $\varphi_{TH}(A)$ ) by a vector is trivially obtained by applying a black box for multiplication of  $A$  by a vector, twice, and performing  $O(n)$  additional operations in  $\mathbb{F}$ . The following, is therefore, immediate from Lemma 3.1.

**Theorem 4.1.** *One can check whether a matrix  $A \in \mathbb{F}^{n \times n}$  is Toeplitz-like, Hankel-like, or Toeplitz+Hankel-like, with displacement rank at most  $k$ , and return a representation of the operator matrix of  $A$  allowing a superfast algorithm to be applied to  $A$  if this is the case.*



The cost to check for these properties, produce and return the above representation of the operator matrix, and verify it — and the probabilities and types of failures of these processes — are as described in Lemma 3.1 for the detection, certification and verification of a matrix with low rank.

## 5 Additive Conditioners

Recall that the *invariant factors* of a matrix  $A \in \mathbb{F}^{n \times n}$  are monic polynomials

$$\varphi_1, \varphi_2, \dots, \varphi_m \in \mathbb{F}[z],$$

each with positive degree, such that  $\varphi_i$  is divisible by  $\varphi_{i+1}$  for  $1 \leq i \leq m-1$  and such that  $A$  is similar to a block diagonal matrix with the companion matrices of the polynomials  $\varphi_1, \varphi_2, \dots, \varphi_m$  as its blocks. In this case  $\varphi_1$  is the minimal polynomial of  $A$ . An invariant factor  $\varphi_i$  is a **nontrivial invariant factor** if  $\varphi_i \neq z$  — for its companion matrix is different from the  $1 \times 1$  zero matrix in this case. Additional (trivial) “invariant factors”  $\varphi_i = 1$  will occasionally be added, below, for  $m+1 \leq i \leq n$ , to simplify technical statements.

The number of invariant factors divisible by  $z^2$  is of interest because this is the same as the number of “nontrivial nilpotent blocks” (companion matrices of polynomials  $z^j$  for  $j \geq 2$ ) in a Jordan normal form for  $A$ .

Techniques of Villard [8] that were developed during the study of a black box algorithm for the Frobenius normal form lead to an efficient interactive protocol to bound the number of nontrivial nilpotent blocks. In combination with a recent protocol of Dumas, Kaltöfen, Thomé and Villard [2] for the certification of the minimal polynomial of a matrix, these lead to an efficient protocol to bound the number of nontrivial invariant factors of a matrix as well.

In particular, the following result of Villard [8, Lemma 1] is of use here.

**Theorem 5.1** (Villard [8]). *Let  $A, B \in \mathbb{F}^{n \times n}$  such that the rank of  $B$  is at most  $k$ . If  $s_1, s_2, \dots, s_n$  are the invariant factors of  $A$  and  $\sigma_1, \sigma_2, \dots, \sigma_n$  are the invariant factors of  $A + B$  then  $s_i$  is divisible by  $\sigma_{i+k}$  in  $\mathbb{F}[z]$  for  $1 \leq i \leq n - k$ .*

Villard also provided a result — [8, Theorem 2] — which is of use to confirm that  $A$  does not have  $k$  or more nontrivial nilpotent blocks, or nontrivial invariant factors, when  $\mathbb{F}$  is sufficiently large. The following result complements Villard’s result by allowing this to be checked for, when  $k$  is small, and when  $\mathbb{F}$  is a very small finite field — the case where “preconditioning” is generally most complicated and expensive, so that the assurance that preconditioning can be avoided might be of greatest interest.

**Theorem 5.2.** *Let  $A \in \mathbb{F}^{n \times n}$  where  $\mathbb{F}$  is a finite field with size  $q$ . Let  $B = V \cdot U$  where  $U \in \mathbb{F}^{k \times n}$ ,  $V \in \mathbb{F}^{n \times k}$ , and the entries of  $U$  and  $V$  are selected uniformly and independently from  $\mathbb{F}$ .*

- (a) If  $A$  has at most  $k$  nontrivial nilpotent blocks then the minimal polynomial of  $A + B$  is not divisible by  $z^2$  with probability at least

$$\rho_1(q) = \frac{(q^2 - 2)(q^2 - q - 1)(q - 1)}{q^4(q + 1)} = 1 - \frac{3q^4 + 2q^3 - 5q^2 + 2}{q^5 + q^4} \geq 1 - 3q^{-1}.$$

- (b) If  $A$  has at most  $k$  nontrivial invariant factors and  $f \in \mathbb{F}[z]$  is an irreducible polynomial with degree  $d$  such that  $f \neq z$ , then the probability that  $f$  does not divide the minimal polynomial of  $A + B$  is at least

$$\begin{aligned} \rho_2(q, d) &= \frac{(q^{4d} - 2)(q^{2d} - q^d - 1)}{q^{3d}(q^{3d} + q^{2d} + q^d + 1)} \\ &= 1 - \frac{2q^{5d} + 2q^{4d} + q^{3d} + 2q^{2d} - 2q^d - 2}{q^{3d}(q^{3d} + q^{2d} + q^d + 1)} \geq 1 - 2q^{-d}. \end{aligned}$$

The proof of this result is, regrettably, rather long, but is also reasonably straightforward: Basic linear algebra and probability theory suffice to establish the above result.

**Lemma 5.3.** *Let  $A \in \mathbb{F}^{n \times n}$ .*

- (a) *If  $A$  has rank  $r < n$  and the entries of vectors  $u \in \mathbb{F}^{1 \times n}$  and  $v \in \mathbb{F}^{n \times 1}$  are chosen uniformly and independently from  $\mathbb{F}$  then  $A + v \cdot u$  has rank  $r + 1$  with probability  $(1 - |\mathbb{F}|^{-(n-r)})^2$ , rank  $r - 1$  with probability at most  $|\mathbb{F}|^{-2(n-r)}$ , and rank  $r$ , otherwise.*
- (b) *Let  $\ell$  be a positive integer. Suppose that  $A \in \mathbb{F}^{n \times n}$  is nonsingular,  $v \in \mathbb{F}^{n \times 1}$ , and that  $R \in \mathbb{F}^{\ell \times n}$ . Then either*
- i.  *$A + v \cdot u \cdot R$  is nonsingular for every vector  $u \in \mathbb{F}^{1 \times \ell}$ , or*
  - ii. *if the entries of a vector  $u \in \mathbb{F}^{1 \times \ell}$  are chosen uniformly and independently from  $\mathbb{F}$  (and independently from the entries of  $A$ ,  $v$  and  $R$ ) then  $A + v \cdot u \cdot R$  is nonsingular with probability  $1 - |\mathbb{F}|^{-1}$ .*
- (c) *If  $A$  is nonsingular and  $v = 0 \in \mathbb{F}^{n \times 1}$  then  $A + v \cdot u$  is nonsingular, as well, for every vector  $u \in \mathbb{F}^{1 \times n}$ . If  $v$  is a nonzero vector in  $\mathbb{F}^{n \times 1}$  and the entries of  $u \in \mathbb{F}^{1 \times n}$  are chosen uniformly and independently from  $\mathbb{F}$  then  $A + v \cdot u$  is nonsingular with probability  $1 - |\mathbb{F}|^{-1}$ .*

*Proof.* Suppose first that  $A$  has rank  $r < n$  and the entries of vectors  $u \in \mathbb{F}^{1 \times n}$  and  $v \in \mathbb{F}^{n \times 1}$  are chosen uniformly and independently from  $\mathbb{F}$ .

Since the column space of  $A$  includes  $|\mathbb{F}|^r$  vectors,  $v$  is not in the column space of  $A$  with probability  $1 - |\mathbb{F}|^{-(n-r)}$ . This is a necessary condition for the rank of  $A + v \cdot u$  to exceed that of  $A$ , since the column space of  $A + v \cdot u$  is a subspace of the column space of  $A$ , otherwise.

With that noted, suppose that  $v_1$  is not in the column space of  $A$ .

Let  $w_1, w_2, \dots, w_n \in \mathbb{F}^{n \times 1}$  be the columns of the matrix  $A \in \mathbb{F}^{n \times n}$  being considered. Permuting columns as needed we may assume without loss of generality that the first  $r$  columns of  $A$ ,  $w_1, w_2, \dots, w_r$ , are linearly independent.

Let  $\mu_1, \mu_2, \dots, \mu_n$  be the entries of the vector  $u \in \mathbb{F}^{1 \times n}$ . Since columns  $w_1, w_2, \dots, w_r$  are linearly independent, and  $v$  is not in the column space of  $A$ , it is easily checked that the first  $r$  columns

$$w_1 + \mu_1 v, w_2 + \mu_2 v, \dots, w_r + \mu_r v$$

of the matrix  $A + v \cdot u$  must be linearly independent as well — so that the rank of  $A + v \cdot u$  is at least  $r$ . On the other hand, the column space of this matrix is a subspace of the space spanned by  $w_1, w_2, \dots, w_r, v$ , so that the rank of this matrix is also at most  $r + 1$ .

Consider any choice of the first  $r$  entries,  $\mu_1, \mu_2, \dots, \mu_r$ , of  $u$ , and let  $i$  be an integer such that  $r + 1 \leq i \leq n$ . Since  $A$  has rank  $r$  the  $i^{\text{th}}$  column  $w_i$  of  $A$  must be a linear combination of the first  $r$  columns, so that there exist elements  $\alpha_1, \alpha_2, \dots, \alpha_r \in \mathbb{F}$  such that

$$w_i = \sum_{j=1}^r \alpha_j w_j.$$

Now — again, since  $v$  is not in the column space of  $A$  — it is easily checked that the  $i^{\text{th}}$  column  $w_i + \mu_i v$  of the matrix  $A + v \cdot u$  is only a linear combination of the first  $r$  columns of this matrix if

$$w_i + \mu_i v = \sum_{j=1}^r \alpha_j (w_j + \mu_j v)$$

as well — for the same values  $\alpha_1, \alpha_2, \dots, \alpha_r \in \mathbb{F}$  as above. In this case one can see — by considering the multipliers for  $v$  in the above equation — that it must also be true that

$$\mu_i = \sum_{j=1}^r \alpha_j \mu_j$$

so that there is only one choice of  $\mu_i$  for which this condition holds. Since the values  $\mu_{r+1}, \mu_{r+2}, \dots, \mu_n$  are chosen uniformly and independently from  $\mathbb{F}$ , it now follows that  $A + v \cdot u$  has rank  $r$  (instead of  $r + 1$ ) with probability  $|\mathbb{F}|^{-(n-r)}$  if  $v_1$  is not in the column space of  $A$ . Since the entries of  $u$  and  $v$  are chosen uniformly and independently, it follows that  $A + v \cdot u$  has rank  $r + 1$  with probability  $(1 - |\mathbb{F}|^{-(n-r)})^2$ , as claimed.

As noted above, the rank of  $A + v \cdot u$  can only be  $r - 1$  if  $v$  is in the column space of  $A$ , and the probability of this is  $|\mathbb{F}|^{-(n-r)}$ . Virtually the same argument establishes that the rank of  $A + v \cdot u$  can only be  $r - 1$  if  $u$  is in the row space of  $A$ , as well, and the probability of this is also  $|\mathbb{F}|^{-(n-r)}$ . Since the entries of  $u$  and  $v$  are chosen independently, the probability that  $A + v \cdot u$  has rank  $r - 1$  is at most  $|\mathbb{F}|^{-2(n-r)}$ .

Finally, since the ranks of  $A$  and  $A + v \cdot u$  can differ by at most one, the rank of  $A + v \cdot u$  is in  $\{r - 1, r, r + 1\}$ , as required to complete the proof of part (a) of the claim.

Suppose next that  $A$  is nonsingular,  $v \in \mathbb{F}^{n \times 1}$ ,  $\ell$  is a positive integer, and  $R \in \mathbb{F}^{\ell \times n}$ .

If  $v = 0$ , then  $A + v \cdot u \cdot R = A$  for every vector  $u \in \mathbb{F}^{1 \times \ell}$ , so that  $A + v \cdot u \cdot R$  is certainly nonsingular as well, and case (i), mentioned in the claim, holds.

Otherwise  $A + v \cdot u \cdot R$  is singular if and only if there is a nonzero vector  $x \in \mathbb{F}^{n \times 1}$  such that  $(A + v \cdot u \cdot R)x = Ax + v \cdot (u \cdot R \cdot x) = 0$ . In this case  $Ax$  is a nonzero scalar multiple of  $v$ . Now, since  $(A + v \cdot u \cdot R)x = 0$  if and only if  $(A + v \cdot R \cdot u)(\alpha x) = 0$  for any nonzero  $\alpha \in \mathbb{F}$ , it suffices to consider the unique nonzero vector  $x = -A^{-1}v$  — in which case  $(A + v \cdot u \cdot R)x = 0$  if and only if  $u \cdot (R \cdot x) = 1$ .

If  $R \cdot x = 0 \in \mathbb{F}^{\ell \times 1}$ , then  $(A + v \cdot u \cdot R)x = Ax = -v \neq 0$  for every vector  $u \in \mathbb{F}^{1 \times \ell}$ , and case (i) holds once again.

Suppose, instead, that  $R \cdot x$  is a nonzero vector in  $\mathbb{F}^{\ell \times 1}$  and that the entries of  $u \in \mathbb{F}^{1 \times \ell}$  are chosen uniformly and independently from  $\mathbb{F}$  (and independently of the entries of  $A$ ,  $v$ , and  $R$ ). Consider an integer  $i$  such that  $1 \leq i \leq \ell$  and the  $i^{\text{th}}$  entry of  $R \cdot x$  is nonzero. After all other entries of  $u$  have been selected there is exactly one choice of the  $i^{\text{th}}$  entry of  $u$  such that  $u \cdot R \cdot x = 1$ . Thus  $(A - v \cdot R \cdot u)x = 0$  with probability at most  $|\mathbb{F}|^{-1}$ , so that case (ii) holds — establishing part (b) of the claim.

Part (c) of the claim is a trivial consequence of part (b), obtained by setting  $\ell = n$  and setting  $R$  to be the identity matrix  $I_n \in \mathbb{F}^{n \times n}$ .  $\square$

**Lemma 5.4.** *Let  $A \in \mathbb{F}^{n \times n}$  be a matrix with rank  $r$  for a nonnegative integer  $r$ , and let  $k$  be a positive integer such that  $k \geq n - r$ .*

- (a) *If  $k = n - r$ , and the entries of matrices  $U \in \mathbb{F}^{k \times n}$  and  $V \in \mathbb{F}^{n \times k}$  are chosen uniformly and independently from  $\mathbb{F}$ , then  $A + V \cdot U$  is nonsingular with probability at least*

$$\left(1 - \frac{1}{|\mathbb{F}|^2 - 1}\right) (1 - |\mathbb{F}|^{-1})^2 = \frac{(|\mathbb{F}|^2 - 2)(|\mathbb{F}| - 1)}{|\mathbb{F}|^2(|\mathbb{F}| + 1)} \geq 1 - 2|\mathbb{F}|^{-1}.$$

- (b) *Let  $\ell$  be a positive integer. Suppose that  $A \in \mathbb{F}^{n \times n}$  is nonsingular,  $V \in \mathbb{F}^{n \times k}$ , and  $R \in \mathbb{F}^{\ell \times n}$ . If the entries of  $U \in \mathbb{F}^{k \times \ell}$  are chosen uniformly and independently from  $\mathbb{F}$  (and independently from the entries of  $A$ ,  $V$  and  $R$ ) then  $A + V \cdot U \cdot R$  is nonsingular with probability at least*

$$\frac{|\mathbb{F}| - 1}{|\mathbb{F}|} = 1 - |\mathbb{F}|^{-1}$$

*if  $k = 1$ , and with probability at least*

$$\left(1 - \frac{|\mathbb{F}|^{-1}}{|\mathbb{F}| - 1}\right) \cdot (1 - |\mathbb{F}|^{-1}) = 1 - \frac{|\mathbb{F}| + 1}{|\mathbb{F}|^2} > 1 - 2|\mathbb{F}|^{-1}$$

*when  $k \geq 2$ .*

- (c) If  $A$  is nonsingular,  $V \in \mathbb{F}^{n \times k}$ , and the entries of a matrix  $U \in \mathbb{F}^{k \times n}$  are chosen uniformly and independently from  $\mathbb{F}$  (and independently from the entries of  $A$  and  $V$ ), then the matrix  $A + V \cdot U$  is nonsingular with probability at least

$$\frac{|\mathbb{F}| - 1}{|\mathbb{F}|} = 1 - |\mathbb{F}|^{-1}$$

if  $k = 1$ , and with probability at least

$$\left(1 - \frac{|\mathbb{F}|^{-1}}{|\mathbb{F}| - 1}\right) \cdot (1 - |\mathbb{F}|^{-1}) = 1 - \frac{|\mathbb{F}| + 1}{|\mathbb{F}|^2} > 1 - 2|\mathbb{F}|^{-1}$$

when  $k \geq 2$ .

- (d) If  $n - r < k$  and  $A$  has rank  $r < n$ , and the entries of matrices  $U \in \mathbb{F}^{k \times n}$  and  $V \in \mathbb{F}^{n \times k}$  are chosen uniformly and independently from  $\mathbb{F}$ , then  $A + V \cdot U$  is nonsingular with probability at least

$$\frac{(|\mathbb{F}|^4 - 2)(|\mathbb{F}|^2 - |\mathbb{F}| - 1)}{|\mathbb{F}|^3(|\mathbb{F}|^3 + |\mathbb{F}|^2 + |\mathbb{F}| + 1)} = 1 - \frac{2|\mathbb{F}|^5 + 2|\mathbb{F}|^4 + |\mathbb{F}|^3 + 2|\mathbb{F}|^2 - 2|\mathbb{F}| - 2}{|\mathbb{F}|^3(|\mathbb{F}|^3 + |\mathbb{F}|^2 + |\mathbb{F}| + 1)} \geq 1 - 2|\mathbb{F}|^{-1}.$$

*Proof.* Suppose first that  $A \in \mathbb{F}^{n \times n}$  is an arbitrarily chosen matrix with rank  $r$  such that  $k \geq n - r$ . Note that if a matrix  $U \in \mathbb{F}^{k \times n}$  has rows  $u_1, u_2, \dots, u_k \in \mathbb{F}^{1 \times n}$  (from top to bottom) and  $V \in \mathbb{F}^{n \times k}$  has columns  $v_1, v_2, \dots, v_k \in \mathbb{F}^{n \times 1}$  (from left to right) then

$$A + V \cdot U = A + \sum_{i=1}^k v_i \cdot u_i.$$

With that noted — supposing, as above, that the entries of  $U \in \mathbb{F}^{k \times n}$  and  $V \in \mathbb{F}^{n \times k}$  are chosen uniformly and independently from  $\mathbb{F}$  — let  $\rho_{k,r}$  be the probability that there exist integers  $j_1, j_2, \dots, j_{n-r}$  such that

$$1 \leq j_1 < j_2 < \dots < j_{n-r} \leq k$$

and such that the matrix

$$A + \sum_{h=1}^{n-r} v_{j_h} \cdot u_{j_h}$$

is nonsingular. It trivially follows that  $\rho_{k,n} = 1$  for every integer  $k$  such that  $0 \leq k \leq n$ .

Suppose next that  $r < n$ . Since the entries of  $U \in \mathbb{F}^{k \times n}$  and  $V \in \mathbb{F}^{n \times k}$  are chosen uniformly and independently from  $\mathbb{F}$ , it follows by a straightforward application of part (a) of Lemma 5.3 that the probability that  $A + v_i \cdot u_i$  has rank at most  $r$ , for every integer  $i$  such that  $1 \leq i \leq k$ , is  $(2|\mathbb{F}|^{r-n} - |\mathbb{F}|^{2(r-n)})^k$ .

Suppose, instead, that there exists an integer  $i$  such that  $1 \leq i \leq k$  and  $A + v_i \cdot u_i$  has rank  $r + 1$ . Permuting the rows of  $U$  and columns of  $V$  as needed — without changing the distributions used to generate these matrices — we may assume without loss of generality that  $i = k$ . Now let

$$\hat{A} = A + v_k \cdot u_k \in \mathbb{F}^{n \times n},$$

a matrix with rank  $r + 1$ . Since the entries of the vectors  $u_1, u_2, \dots, u_{k-1} \in \mathbb{F}^{1 \times n}$  and the vectors  $v_1, v_2, \dots, v_{k-1} \in \mathbb{F}^{n \times 1}$  are chosen uniformly and independently from  $\mathbb{F}$  — and independently of either the entries of  $u_k \in \mathbb{F}^{1 \times n}$  and  $v_k \in \mathbb{F}^{n \times 1}$ , or of the entries of the above matrix  $\hat{A}$  — one can now consider  $\hat{A}$  instead of  $A$  to conclude that

$$\rho_{k,r} = \left( 1 - \left( 2|\mathbb{F}|^{r-n} - |\mathbb{F}|^{2(r-n)} \right)^k \right) \cdot \rho_{k-1,r+1}.$$

It follows, by induction on  $n - r$ , that

$$\begin{aligned} \rho_{k,r} &= \left( 1 - \left( 2|\mathbb{F}|^{r-n} - |\mathbb{F}|^{2(r-n)} \right)^k \right) \cdot \rho_{k-1,r+1} \\ &= \prod_{i=0}^{n-r-1} \left( 1 - \left( 2|\mathbb{F}|^{r-n+i} - |\mathbb{F}|^{2(r-n+i)} \right)^{k-i} \right). \end{aligned}$$

If  $0 \leq i \leq n - r - 2$  then

$$\begin{aligned} 1 - \left( 2|\mathbb{F}|^{r-n+i} - |\mathbb{F}|^{2(r-n+i)} \right)^{k-i} &= 1 - |\mathbb{F}|^{(r-n+i)(k-i)} (2 - |\mathbb{F}|^{r-n+i})^{k-i} \\ &\geq 1 - (2|\mathbb{F}|^{r-n+i})^{k-i} \\ &\geq 1 - (|\mathbb{F}|^{k-i})^{r-n+i+1}. \end{aligned}$$

If  $i = n - r - 1$  then

$$1 - \left( 2|\mathbb{F}|^{r-n+i} - |\mathbb{F}|^{2(r-n+i)} \right)^{k-i} = 1 - (2|\mathbb{F}|^{-1} - |\mathbb{F}|^{-2})^{k-(n-r)+1}.$$

Thus

$$\begin{aligned} \rho_{k,r} &\geq \left( \prod_{i=0}^{n-r-2} 1 - (|\mathbb{F}|^{k-i})^{r-n+i+1} \right) \cdot (1 - (2|\mathbb{F}|^{-1} - |\mathbb{F}|^{-2})^{k-(n-r)+1}) \\ &= \left( \prod_{i=0}^{n-r-2} 1 - (|\mathbb{F}|^{i-k})^{n-r-i-1} \right) \cdot (1 - (2|\mathbb{F}|^{-1} - |\mathbb{F}|^{-2})^{k-(n-r)+1}) \\ &\geq \left( \prod_{i=0}^{n-r-2} 1 - (|\mathbb{F}|^{(n-r)-k-2})^{n-r-i-1} \right) \cdot (1 - (2|\mathbb{F}|^{-1} - |\mathbb{F}|^{-2})^{k-(n-r)+1}) \\ &\quad (\text{since } (n-r) - k - 2 \geq i - k \text{ when } 0 \leq i \leq n - r - 2) \end{aligned}$$

$$\begin{aligned}
&\geq \left(1 - \sum_{i=0}^{n-r-2} (|\mathbf{F}|^{(n-r)-k-2})^{r-n-i-1}\right) \cdot (1 - (2|\mathbf{F}|^{-1} - |\mathbf{F}|^{-2})^{k-(n-r)+1}) \\
&\geq \left(1 - \sum_{j \geq 1} (|\mathbf{F}|^{(n-r)-k-2})^j\right) \cdot (1 - (2|\mathbf{F}|^{-1} - |\mathbf{F}|^{-2})^{k-(n-r)+1}) \\
&= \left(1 - \frac{1}{|\mathbf{F}|^{k-(n-r)+2} - 1}\right) \cdot (1 - (2|\mathbf{F}|^{-1} - |\mathbf{F}|^{-2})^{k-(n-r)+1}) \tag{5.1}
\end{aligned}$$

Suppose, now, that  $k = n - r$ , so that  $r < n$  since  $k$  is a positive integer. Then

$$1 - (2|\mathbf{F}|^{-1} - |\mathbf{F}|^{-2})^{k-(n-r)+1} = (1 - |\mathbf{F}|^{-1})^2.$$

It is easily checked (setting  $k = n - r$ ) that this establishes part (a) of the claim.

Part (b) follows by a similar argument: Suppose, now, that  $A$  is nonsingular, let  $k \geq 0$ ,  $V \in \mathbf{F}^{n \times k}$ ,  $\ell \geq 1$ ,  $R \in \mathbf{F}^{\ell \times n}$ , and suppose that the entries of  $U \in \mathbf{F}^{k \times \ell}$  are chosen uniformly and independently from  $\mathbf{F}$  (and independently of the entries of  $A$ ,  $V$  and  $R$ ). Let  $\mu_k$  be the probability that the matrix  $A + V \cdot U \cdot R$  is also nonsingular. It trivially follows that  $\mu_0 = 1$ .

Suppose next that  $k \geq 1$ . The probability that  $A + V \cdot U \cdot R$  is nonsingular can be under-approximated by the probability that both this is the case and there exists an integer  $i$  such that  $1 \leq i \leq k$  and  $A + v_i \cdot u_i \cdot R$  is nonsingular where  $u_1, u_2, \dots, u_k \in \mathbf{F}^{1 \times \ell}$  are the rows of  $U$  (from top to bottom) and  $v_1, v_2, \dots, v_k \in \mathbf{F}^{n \times 1}$  are the columns of  $v$  (from left to right).

Since the entries of  $U \in \mathbf{F}^{k \times \ell}$  are chosen uniformly and independently from  $\mathbf{F}$  (and independently of the entries of  $A$ ,  $V$  and  $R$ ), it follows by a straightforward application of part (b) of Lemma 5.3 that  $A + v_i \cdot u_i \cdot R$  is singular, for every integer  $i$  such that  $1 \leq i \leq k$ , with probability at most  $|\mathbf{F}|^{-k}$ .

Suppose now that  $A + v_i \cdot u_i \cdot R$  is nonsingular for at least one integer  $i$  such that  $1 \leq i \leq k$ . Once again, permuting the rows of  $U$  and columns of  $V$  as needed — without changing the distribution to generate these matrices — we may assume without loss of generality that  $i = k$ . Let

$$\widehat{A} = A + v_k \cdot u_k \cdot R,$$

a nonsingular matrix in  $\mathbf{F}^{n \times n}$ . Since the entries of the vectors  $u_1, u_2, \dots, u_{k-1} \in \mathbf{F}^{1 \times \ell}$  and vectors  $v_1, v_2, \dots, v_{k-1} \in \mathbf{F}^{n \times 1}$  are chosen uniformly and independently from  $\mathbf{F}$  — and independently of either the entries of the vectors  $u_k \in \mathbf{F}^{1 \times n}$  and  $v_k \in \mathbf{F}^{n \times 1}$  or the entries of the above matrix  $\widehat{A}$  — we now have that

$$\mu_k \geq (1 - |\mathbf{F}|^{-k}) \cdot \mu_{k-1}.$$

Thus  $\mu_1 \geq 1 - |\mathbf{F}|^{-1}$  and it is easily established by induction on  $k$  that if  $k \geq 2$  then

$$\mu_k \geq \prod_{i=0}^{k-1} (1 - |\mathbf{F}|^{-i-k})$$

$$\begin{aligned}
&\geq \left( \prod_{j \geq 2} (1 - |\mathbf{F}|^{-j}) \right) \cdot (1 - |\mathbf{F}|^{-1}) \\
&\geq \left( 1 - \sum_{j \geq 2} |\mathbf{F}|^{-j} \right) \cdot (1 - |\mathbf{F}|^{-1}) \\
&+ \left( 1 - \frac{|\mathbf{F}|^{-1}}{|\mathbf{F}| - 1} \right) \cdot (1 - |\mathbf{F}|^{-1}) \\
&= 1 - \frac{|\mathbf{F}| + 1}{|\mathbf{F}|^2} \\
&\geq 1 - 2|\mathbf{F}|^{-1},
\end{aligned}$$

as needed to establish part (b) of the claim.

Part (c) of the claim is a trivial corollary of part (b), obtained by setting  $\ell = n$  and setting  $R$  to be the identity matrix  $I_n \in \mathbb{F}^{n \times n}$ .

Essentially the same argument (with  $\ell = n$  and  $R = I_n$ ) establishes, for  $k > n - r$ , that if  $A \in \mathbb{F}^{n \times n}$  has rank  $r$ , then the conditional probability that  $A + V \cdot U$  is nonsingular, given that there exist integers  $j_1, j_2, \dots, j_{n-r}$  such that

$$1 \leq j_1 < j_2 < \dots < j_{n-r} \leq k$$

and the matrix

$$A + \sum_{h=1}^{n-r} v_{j_h} \cdot u_{j_h}$$

is nonsingular, is at least

$$1 - |\mathbf{F}|^{-1} \tag{5.2}$$

if  $k = n - r + 1$ , and at least

$$1 - \frac{|\mathbf{F}| + 1}{|\mathbf{F}|^2} \geq 1 - 2|\mathbf{F}|^{-1} \tag{5.3}$$

if  $k \geq n - r + 2$ .

One can now under-approximate the probability that  $A + V \cdot U$  is nonsingular, when  $A$  has rank  $r < n$  and  $k > n - r$ , by the probability that both this is the case and there exist integers  $j_1, j_2, \dots, j_{n-r}$  such that

$$1 \leq j_1 < j_2 < \dots < j_{n-r} \leq k$$

and the matrix  $A + \sum_{h=1}^{n-r} v_{j_h} \cdot u_{j_h}$  is nonsingular. It follows by the bounds at lines (5.1)

and (5.2) that this is at least  $f_1(|\mathbf{F}|)$ , where

$$f_1(z) = \left( 1 - \frac{1}{z^3 - 1} \right) \cdot (1 - (2z^{-1} - z^{-2})^2) \cdot (1 - z^{-1})$$



$$= \frac{(z^3 - 2)}{z(z^2 + z + 1)} \cdot (1 - (2z^{-1} - z^{-2})^2)$$

if  $k = n - r + 1$ , and — by the inequalities at lines (5.1) and (5.3) — at least  $f_2(|F|)$ , where

$$\begin{aligned} & \left(1 - \frac{1}{z^4 - 1}\right) \cdot (1 - (2z^{-1} - z^{-2})^3) \cdot \left(\frac{z^2 - z - 1}{z^2}\right) \\ & \geq \left(1 - \frac{1}{z^4 - 1}\right) \cdot (1 - z^{-1}) \cdot \left(\frac{z^2 - z - 1}{z^2}\right) \\ & = \frac{(z^4 - 2)(z^2 - z - 1)}{z^3(z^3 + z^2 + z + 1)} = f_2(z) \end{aligned}$$

if  $k \geq n - r + 2$ .

Suppose first that  $|F| = 2$ . Then  $f_1(|F|) = \frac{3}{16}$  and  $f_2(|F|) = \frac{7}{60}$ , so that  $f_2(|F|) \leq f_1(|F|)$  in this case.

On the other hand, if  $z = |F| \geq 3$  then

$$\begin{aligned} f_1(z) &= \frac{(z^3 - 2)}{z(z^2 + z + 1)} \cdot (1 - (2z^{-1} - z^{-2})^2) \\ &\geq \frac{(z^3 - 2)}{z(z^2 + z + 1)} \cdot (1 - z^{-1}) \\ &= \frac{(z^3 - 2)(z - 1)}{z^2(z^2 + z + 1)} = \hat{f}_1(z). \end{aligned}$$

Supposing, again, that  $z \geq 3$ , consider the polynomial

$$\begin{aligned} F(z) &= z^3(z^2 + z + 1)(z^3 + z^2 + z + 1) \cdot (\hat{f}_1(z) - f_2(z)) \\ &= z^6 + 2z^4 - 2z^2 - 2z - 2 \in \mathbb{Z}[z]. \end{aligned}$$

Notice that the leading coefficient, 6, of

$$F'(z) = 6z^5 + 8z^3 - 4z - 2,$$

is equal to the sum of the absolute values of all negative coefficients. Since this polynomial includes a second term with a positive coefficient,  $F'(z) > 0$  whenever  $z \geq 1$ . It therefore suffices to confirm that  $F(3) = 865 > 0$  to confirm that  $\hat{f}_1(|F|) \geq f_2(|F|)$  whenever  $|F| \geq 3$ . Thus the desired probability is always at least  $f_2(|F|)$ , as needed to establish part (d) of the claim.  $\square$

**Lemma 5.5.** *If  $n$  and  $t$  are positive integers, and  $A \in \mathbb{F}^{n \times n}$  is a matrix with rank  $n - t$  such that the characteristic polynomial of  $A$  is  $z^t \varphi(z)$  for a polynomial  $\varphi \in \mathbb{F}[z]$  such that  $\varphi(0) \neq 0$ , then the minimal polynomial of  $A$  is not divisible by  $z^2$ .*

*Proof.* Suppose, to the contrary, that the minimal polynomial of  $A$  is divisible by  $z^2$ . Then, since  $A$  has rank  $n - t$ , the  $i^{\text{th}}$  invariant factor must be divisible by  $z^{j_i}$  for an integer  $j_i$  and for  $1 \leq i \leq t$ , where

$$j_1 \geq j_2 \geq \cdots \geq j_t \geq 1.$$

Furthermore  $j_1 \geq 2$  since the minimal polynomial of  $A$  is divisible by  $z^2$ .

Since the characteristic polynomial of  $A$  is the product of the invariant factors of  $A$ , it follows that the characteristic polynomial of  $A$  must be divisible by  $\prod_{i=1}^t z^{j_i} = z^{\sum_{i=1}^t j_i}$  and, since  $\sum_{i=1}^t j_i \geq t + 1$ , the characteristic polynomial of  $A$  cannot be as described in the claim.  $\square$

**Theorem 5.6.** *Let  $k$  be a positive integer and let  $A \in \mathbb{F}^{n \times n}$ , for a positive integer  $n$ , such that  $A$  has at most  $k$  nontrivial nilpotent blocks. If the entries of  $U \in \mathbb{F}^{k \times n}$  and  $V \in \mathbb{F}^{n \times k}$  are chosen uniformly and independently from  $\mathbb{F}$  then the probability that the minimal polynomial of  $A + V \cdot U$  is not divisible by  $z^2$  is at least*

$$\frac{(|\mathbb{F}|^2 - 2)(|\mathbb{F}|^2 - |\mathbb{F}| - 1)(|\mathbb{F}| - 1)}{|\mathbb{F}|^4(|\mathbb{F}| + 1)} = 1 - \frac{3|\mathbb{F}|^4 + 2|\mathbb{F}|^3 - 5|\mathbb{F}|^2 + 2}{|\mathbb{F}|^5 + |\mathbb{F}|^4} \geq 1 - 3|\mathbb{F}|^{-1} \quad (5.4)$$

*Proof.* Let  $\ell$  be the number of nontrivial nilpotent blocks of  $A$ , so that  $\ell \leq k$ . Suppose that  $A$  has rank  $r = n - t$  for an integer  $t$ . Then  $\ell \leq t \leq n$ . The cases that  $t \leq k$  and  $t > k$  are considered, separately, below.

Suppose first that  $t \leq k$ . It follows by parts (a) and (d) of Lemma 5.4 that  $A + V \cdot U$  is nonsingular with probability at least  $\min(f_1(|\mathbb{F}|), f_2(|\mathbb{F}|))$ , where

$$f_1(z) = \frac{(z^2 - 2)(z - 1)}{z^2(z + 1)}$$

and

$$f_2(z) = \frac{(z^4 - 2)(z^2 - z - 1)}{z^3(z^3 + z^2 + z + 1)}.$$

The minimal polynomial of  $A + V \cdot U$  cannot be divisible by either  $z$  or  $z^2$  if this is the case.

Suppose next that  $t > k$ , so that the rank of  $A$  is  $n - t < n - k$ . In this case — since the number of nilpotent blocks with size one in a (rational) Jordan normal form for  $A$  is  $t - \ell \geq t - k$  —  $A$  is similar to a matrix

$$\tilde{A} = \begin{bmatrix} \hat{A} & 0_{(n-t+k) \times (t-k)} \\ 0_{(t-k) \times (n-t+k)} & 0_{(t-k) \times (t-k)} \end{bmatrix}, \quad (5.5)$$

where  $\hat{A} \in \mathbb{F}^{(n-t+k) \times (n-t+k)}$ , so that  $A = X^{-1} \tilde{A} X$  for a nonsingular matrix  $X \in \mathbb{F}^{n \times n}$ .

Now, since  $A$  and  $\tilde{A}$  are similar, these matrices have the same rank, invariant factors, and the same number of nontrivial nilpotent blocks. Furthermore, if the entries of matrices

$U \in \mathbb{F}^{k \times n}$  and  $V \in \mathbb{F}^{n \times k}$  are chosen uniformly and independently from  $\mathbb{F}$  then so are the entries of the matrices  $\tilde{U} = U \cdot X^{-1} \in \mathbb{F}^{k \times n}$ , and  $\tilde{V} = X \cdot V \in \mathbb{F}^{n \times k}$ . Multiplying by  $X$  on the left and by  $X^{-1}$  we may therefore replace  $A$  with  $\tilde{A}$  — effectively assuming without loss of generality that  $A = \tilde{A}$  as shown at line (5.5).

Let  $U_L \in \mathbb{F}^{k \times (n-t+k)}$  and  $U_R \in \mathbb{F}^{k \times (t-k)}$  be the left and right submatrices of  $U$ , and let  $V_T \in \mathbb{F}^{(n-t+k) \times k}$  and  $V_B \in \mathbb{F}^{(t-k) \times k}$  be the top and bottom submatrices of  $V$ , so that

$$A + V \cdot U = \begin{bmatrix} \hat{A} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} V_T \\ V_B \end{bmatrix} \cdot \begin{bmatrix} U_L & U_R \end{bmatrix} = \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix}$$

where

$$A_{1,1} = \hat{A} + V_T \cdot U_L \in \mathbb{F}^{(n-t+k) \times (n-t+k)},$$

$$A_{1,2} = V_T \cdot U_R \in \mathbb{F}^{(n-t+k) \times (t-k)},$$

$$A_{2,1} = V_B \cdot U_L \in \mathbb{F}^{(t-k) \times (n-t+k)},$$

and

$$A_{2,2} = V_B \cdot U_R \in \mathbb{F}^{(t-k) \times (t-k)}.$$

Since the entries of  $U_L$  and  $V_T$  are chosen uniformly and independently from  $\mathbb{F}$ , it follows by part (a) of Lemma 5.4 (replacing  $n$  with  $n - t + k$  and replacing  $r$  with  $n - t$ ) that the matrix  $A_{1,1} = \hat{A} + V_T \cdot U_L$  is nonsingular with probability at least

$$\left(1 - \frac{1}{|\mathbb{F}|^2 - 1}\right) \cdot (1 - |\mathbb{F}|^{-1})^2 = \frac{(|\mathbb{F}|^2 - 2)(|\mathbb{F}| - 1)}{|\mathbb{F}|^2(|\mathbb{F}| + 1)}.$$

With that noted, consider (for the rest of this argument) the case that  $A_{1,1}$  is, indeed, nonsingular.

In this case, since the rank of  $A + V \cdot U$  cannot exceed  $n - t + k$ , it follows that there exist matrices  $Y \in \mathbb{F}^{(n-t+k) \times (t-k)}$  and  $Z \in \mathbb{F}^{(t-k) \times (n-t+k)}$  such that  $A_{1,2} = A_{1,1} \cdot Y$ , and  $A_{2,1} = Z \cdot A_{1,1}$  and — since  $A + V \cdot U$  and  $A_{1,1}$  have the same rank —  $A_{2,2} = Z \cdot A_{1,1} \cdot Y$ . Thus

$$\begin{aligned} A + V \cdot U &= \begin{bmatrix} A_{1,1} & A_{1,1} \cdot Y \\ Z \cdot A_{1,1} & Z \cdot A_{1,1} \cdot Y \end{bmatrix} \\ &= \begin{bmatrix} I_{n-t+k} & 0 \\ Z & I_{k-t} \end{bmatrix} \cdot \begin{bmatrix} A_{1,1} & 0 \\ 0 & 0_{k-t} \end{bmatrix} \cdot \begin{bmatrix} I_{n-t+k} & Y \\ 0 & I_{k-t} \end{bmatrix}. \end{aligned}$$

Since the rightmost matrix shown in the above line is nonsingular,  $A + V \cdot U$  is similar to the matrix

$$\begin{bmatrix} I_{n-t+k} & Y \\ 0 & I_{k-t} \end{bmatrix} \cdot \begin{bmatrix} I_{n-t+k} & 0 \\ Z & I_{k-t} \end{bmatrix} \cdot \begin{bmatrix} A_{1,1} & 0 \\ 0 & 0_{k-t} \end{bmatrix}$$

$$\begin{aligned}
&= \begin{bmatrix} I_{n-t+k} & Y \\ 0 & I_{k-t} \end{bmatrix} \cdot \begin{bmatrix} A_{1,1} & 0 \\ ZA_{1,1} & 0_{k-t} \end{bmatrix} \\
&= \begin{bmatrix} (I_{n-t+k} + YZ) \cdot A_{1,1} & 0 \\ ZA_{1,1} & 0_{k-t} \end{bmatrix}.
\end{aligned}$$

The characteristic polynomial of  $A + UV$  is, therefore, the product of  $z^{k-t}$  and the characteristic polynomial of the matrix  $(I_{n-t+k} + YZ) \cdot A_{1,1}$ . Since  $A_{1,1}$  is nonsingular, the matrix  $(I_{n-t+k} + YZ)A_{1,1}$  is also nonsingular if  $I_{n-t+k} + YZ$  is — and it would then follow by Lemma 5.5 (with  $t$  replaced by  $t - k$ ) that the minimal polynomial of  $A + V \cdot U$  is not divisible by  $z^2$ .

It now suffices to note that, since  $Y = A_{1,1}^{-1} \cdot A_{1,2} = A_{1,1}^{-1} \cdot V_T \cdot U_R$  and  $Z = A_{2,1} \cdot A_{1,1}^{-1} = V_B \cdot U_L \cdot A_{1,1}^{-1}$ ,  $I_{n-t+k} + YZ = I_{n-t+k} + \widehat{V} \cdot \widehat{U} \cdot \widehat{R}$  for the matrices

$$\widehat{V} = A_{1,1}^{-1} V_T \in \mathbb{F}^{(n-t+k) \times k},$$

$$\widehat{U} = U_R \in \mathbb{F}^{k \times (t-k)},$$

and

$$\widehat{R} = V_B \cdot U_L \cdot A_{1,1}^{-1} \in \mathbb{F}^{(t-k) \times (n-t+k)}.$$

Since the entries of  $\widehat{U} = U_R$  are chosen uniformly from  $\mathbb{F}$ , and independently from the entries of  $A$ ,  $U_L$ ,  $V_T$  and  $V_B$ , the entries of  $\widehat{U}$  are also chosen independently from those of  $\widehat{V}$  and  $\widehat{R}$ . Since  $I_{n-t+k}$  is a fixed nonsingular matrix, it follows by part (b) of Lemma 5.4 that the conditional probability that  $I_{n-t+k} + \widehat{V} \cdot \widehat{U} \cdot \widehat{R}$  is nonsingular, given that  $A_{1,1}$  is, is at least

$$\left(1 - \frac{|\mathbb{F}|^{-1}}{|\mathbb{F}| - 1}\right) \cdot (1 - |\mathbb{F}|^{-1}) = \frac{|\mathbb{F}|^2 - |\mathbb{F}| - 1}{|\mathbb{F}|^2},$$

so that the minimal polynomial of  $A + V \cdot U$  is not divisible by  $z^2$ , in this case, with probability at least  $f_3(|\mathbb{F}|)$  where

$$f_3(z) = \frac{(z^2 - 2)(z^2 - z - 1)(z - 1)}{z^4(z + 1)}.$$

Now consider the polynomial

$$F_1(z) = z^3 \cdot (z + 1) \cdot (z^3 + z^2 + z + 1) \cdot (f_1(z) - f_2(z)) = z^4 + z^3 - 2z - 2 \in \mathbb{Z}[z]$$

and its derivative,

$$F_1'(z) = 4z^3 + 3z^2 - 2.$$

Since the leading coefficient, 4, of  $F_1'$  is greater than the sum of the absolute values of the negative coefficients of this polynomial,  $F_1'(z) > 0$  whenever  $z \geq 1$ . Since  $F_1(2) = 18 > 0$ , it follows that  $F_1(z) > 0$  when  $z \geq 2$ , and that  $f_1(z) \geq f_2(z)$  when  $z \geq 2$  as well.

Consider as well the polynomial

$$\begin{aligned} F_2(z) &= z^4 \cdot (z^3 + z^2 + z + 1) \cdot (z + 1) \cdot (f_2(z) - f_3(z)) \\ &= z^7 + z^6 - 3z^5 - 3z^4 - z^3 + z^2 + 4z + 2 \end{aligned}$$

and its derivatives

$$\begin{aligned} F_2'(z) &= 7z^6 + 6z^5 - 15z^4 - 12z^3 - 3z^2 + 2z + 4, \\ F_2''(z) &= 42z^5 + 30z^4 - 60z^3 - 36z^2 - 6z + 2, \end{aligned}$$

and

$$F_2'''(z) = 210z^4 + 120z^3 - 180z^2 - 72z - 6.$$

The leading pair of coefficients of  $F_2'''$  are both positive and their sum, 360, exceeds the sum of the absolute values of the negative coefficients of this polynomial. It follows that  $F_2'''(z) > 0$  when  $z \geq 1$ .

Since  $F_2''(2) = 1190 > 0$ , it now follows that  $F_2''(z) > 0$  when  $z \geq 2$ . Since  $F_2'(2) = 300 > 0$ , it also follows that  $F_2'(z) > 0$  when  $z \geq 2$ . Finally, since  $F_2(2) = 54 > 0$ , it follows that  $F_2(z) > 0$  — and  $f_2(z) \geq f_3(z)$  — whenever  $z \geq 2$  as well.

Thus

$$\begin{aligned} f_3(|F|) &= \frac{(|F|^2 - 2)(|F|^2 - |F| - 1)(|F| - 1)}{|F|^4(|F| + 1)} \\ &= 1 - \frac{3|F|^4 + 2|F|^3 - 5|F|^2 + 2}{|F|^5 + |F|^4} \\ &\geq 1 - \frac{3|F|^4 + 2|F|^3}{|F|^5 + |F|^4} \\ &\geq 1 - 3|F|^{-1} \end{aligned}$$

is a lower bound for the probability that the minimal polynomial of  $A + V \cdot U$  is not divisible by  $z^2$ , in all cases, as needed to establish the claim.  $\square$

The next lemma and theorem generalize parts (a) and (c) of Lemmas 5.3, as well as parts (a), (c) and (d) of 5.4, as needed to establish the second part of Theorem 5.2.

**Lemma 5.7.** *Let  $A \in \mathbb{F}^{n \times n}$  and let  $f \in \mathbb{F}[z]$  be a monic irreducible polynomial with degree  $d$ . Suppose the entries of vectors  $v \in \mathbb{F}^{n \times 1}$  and  $u \in \mathbb{F}^{1 \times n}$  are chosen uniformly and independently from  $\mathbb{F}$ .*

- (a) *If  $\ell$  of the invariant factors of  $A$  are divisible by  $f$ , for a positive integer  $\ell$ , then exactly  $\ell - 1$  of the invariant factors of  $A + v \cdot u$  are divisible by  $f$  with probability at least  $(1 - |F|^{-\ell d})^2$ , and exactly  $\ell + 1$  of the invariant factors of  $A + v \cdot u$  are divisible by  $f$  with probability at most  $|F|^{-2\ell d}$ . Exactly  $\ell$  of the invariant factors of  $A$  are divisible by  $f$ , otherwise.*

(b) *If the minimal polynomial of  $A$  is not divisible by  $f$  then the minimal polynomial of  $A + v \cdot u$  is also not divisible by  $f$  with probability at least  $1 - |\mathbf{F}|^d$ .*

*Proof.* Suppose first that  $d = 1$ ; then  $f = z - \lambda$  for some element  $\lambda$  of  $\mathbf{F}$ . In this case, for  $\ell \geq 1$ , exactly  $\ell$  of the invariant factors of  $A$  (respectively,  $A + v \cdot u$ ) are divisible by  $f$  if and only if  $A - \lambda I_n$  (respectively,  $A - \lambda I_n + v \cdot u$ ) has rank  $r = n - \ell$ . Furthermore, the minimal polynomial of  $A$  (respectively,  $A + v \cdot u$ ) is not divisible by  $f$  if and only if  $A - \lambda I_n$  (respectively,  $A - \lambda I_n + v \cdot u$ ) is nonsingular. The above claims are, therefore, consequences of parts (a) and (c) of Lemma 5.3 in this case.

Suppose next that  $d \geq 2$ . Let  $\mathbf{K} = \mathbf{F}[\lambda] \cong \mathbf{F}[x]/\langle f \rangle$ , where  $\lambda \in \mathbf{K}$  is a root of  $f$  in  $\mathbf{K}$ . Suppose, as well, that

$$f = z^d - \gamma_{d-1}z^{d-1} - \gamma_{d-1}z^{d-2} - \cdots - \gamma_1z - \gamma_0$$

for  $\gamma_{d-1}, \gamma_{d-2}, \dots, \gamma_1, \gamma_0 \in \mathbf{F}$ .

In order to establish part (a) of the claim in this case, recall that every matrix  $A \in \mathbf{F}^{n \times n}$  is similar to a matrix in “rational Jordan form” — that is, a matrix whose blocks are the companion matrices of powers of irreducible polynomials in  $\mathbf{F}[z]$ . In particular,  $A$  is similar to such a matrix, where the first  $\ell$  blocks are the companion matrices of polynomials  $f^{j_1}, f^{j_2}, \dots, f^{j_\ell}$ , for integers  $j_1, j_2, \dots, j_\ell$  such that

$$j_1 \geq j_2 \geq \cdots \geq j_\ell \geq 1,$$

and whose remaining blocks are the companion matrices of polynomials that are relatively prime with  $f$ .

Now, if  $X \in \mathbf{F}^{n \times n}$  is a nonsingular matrix then the entries of  $u \cdot X \in \mathbf{F}^{1 \times n}$  and  $X^{-1} \cdot v \in \mathbf{F}^{n \times 1}$  are chosen uniformly and independently from  $\mathbf{F}$  if the entries of  $u \in \mathbf{F}^{1 \times n}$  and  $v \in \mathbf{F}^{n \times 1}$  are. Applying a similarity transformation we may therefore assume without loss of generality that  $A$  is in rational Jordan form and, in particular, has the form described above.

Consider a uniformly selected vector  $v \in \mathbf{F}^{n \times 1}$  — noting that this can be written as

$$v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_\ell \\ v_{\ell+1} \end{bmatrix} \tag{5.6}$$

where  $v_i \in \mathbf{F}^{d \cdot j_i \times 1}$  for  $1 \leq i \leq \ell$  and where  $v_{\ell+1} \in \mathbf{F}^{m \times 1}$  for  $m = n - d \cdot \sum_{i=1}^{\ell} j_i$ . In this case, the entries of the vectors  $v_1, v_2, \dots, v_{\ell+1}$  are selected uniformly and independently from  $\mathbf{F}$ .

Each of the first  $\ell$  blocks,  $C_{f^{j_i}} - \lambda I_{d \cdot j_i}$  of  $A - \lambda I_n$  (for  $1 \leq i \leq \ell$ ) has nullity one, while the remaining blocks of this matrix are nonsingular.

Let  $i$  be an integer such that  $1 \leq i \leq \ell$  and suppose that

$$v_i = \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_{d \cdot j_i - 1} \end{bmatrix} \in \mathbb{F}^{d \cdot j_i \times 1}.$$

The first  $d \cdot j_i - 1$  columns of  $C_{f^{j_i}} - \lambda I_{d \cdot j_i}$  are linearly independent — indeed, the  $h^{\text{th}}$  column has the nonzero entry  $-\lambda$  in position  $h$ , 1 in position  $h + 1$ , and zeroes everywhere else. The above vector  $v_i$  is therefore in the column space of  $C_{f^{j_i}} - \lambda I_{d \cdot j_i}$  if and only if it is a  $\mathbb{K}$ -linear combination of these columns. Applying Gaussian Elimination (and considering entries of this vector from bottom to top) one can confirm that this is the case if and only if

$$\alpha_{d \cdot j_i - 1} + \alpha_{d \cdot j_i - 2} \lambda^{-1} + \cdots + \alpha_1 \lambda^{-(d \cdot j_i - 2)} + \alpha_0 \cdot \lambda^{-(d \cdot j_i - 1)} = 0,$$

which is the case if and only if  $g(\lambda) = 0$  for the polynomial

$$g = \sum_{h=0}^{d \cdot j_i - 1} \alpha_h z^h \in \mathbb{F}[z].$$

This is the case if and only if  $g$  is divisible by  $f$ . Since the coefficients of  $g$  are chosen uniformly and independently from  $\mathbb{F}$ , the probability of this is at most  $|\mathbb{F}|^{-d}$ .

Since  $A - \lambda I_n$  is a block diagonal matrix with  $C_{f^{j_1}} - \lambda \cdot I_{d \cdot j_1}, C_{f^{j_2}} - \lambda \cdot I_{d \cdot j_2}, \dots, C_{f^{j_\ell}} - \lambda \cdot I_{d \cdot j_\ell}$  as the initial  $\ell$  blocks on its diagonal, and the entries of  $v_1, v_2, \dots, v_{\ell+1}$  (and corresponding components of the vector  $u$ ) are chosen uniformly and independently from  $\mathbb{F}$ , it now follows that  $v$  is in the column space of  $A - \lambda I_n$  with probability at most  $|\mathbb{F}|^{-\ell d}$ .

The matrices  $A$  and  $A^T$  have the same invariant factors, and rational Jordan form. Applying the above argument to  $A^T$ , one can see that the probability that  $u$  is in the row space of  $A - \lambda I_n$  is at most  $|\mathbb{F}|^{-\ell d}$  as well.

As argued in the proof of Lemma 5.3, it is necessary for  $v$  to be in the column space of  $A - \lambda I_n$  and for  $u$  to be in the row space of  $A - \lambda I_n$  in order for the rank of  $A - \lambda I_n + v \cdot u$  to be less than that of  $A - \lambda I_n$ , and the rank of  $A - \lambda I_n + v \cdot u$  cannot be less than  $n - \ell - 1$ . The number of invariant factors of  $A$  therefore divisible by  $f$  is  $\ell + 1$  with probability at most  $|\mathbb{F}|^{-2\ell d}$ . There are never more than  $\ell + 1$  invariant factors of this matrix that are divisible by  $f$ .

Suppose, once again, that  $v$  is not in the column space of  $A - \lambda I_n$ , so that the null space of  $A - \lambda I_n + v \cdot u$  is a subspace of the null space of  $A - \lambda I_n$ . It follows that the null space of  $A - \lambda I_n + v \cdot u$  is a proper subset of the null space of  $A - \lambda I_n$ , so that  $A + v \cdot u$  has at most  $\ell - 1$  invariant factors divisible by  $f$ , if and only if there exists a vector  $x \in \mathbb{F}^{n \times 1}$  such that  $(A - \lambda I_n)x = 0 \neq (A - \lambda I_n + v \cdot u)x = v \cdot (u \cdot x)$ . This is the case if and only if  $u \cdot x \neq 0$ .

If  $A$  is as described above, then it suffices to consider (as  $x$ )  $\ell$  vectors

$$\widehat{x}_1 = \begin{bmatrix} x_1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \widehat{x}_2 = \begin{bmatrix} 0 \\ x_2 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \widehat{x}_3 = \begin{bmatrix} 0 \\ 0 \\ x_3 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \dots, \widehat{x}_\ell = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ x_\ell \\ 0 \end{bmatrix} \quad (5.7)$$

where  $x_i$  is a nonzero vector in  $\mathbb{K}^{d \cdot j_i \times 1}$  such that  $(C_{f^{j_i}} - \lambda I_{d \cdot j_i})x_i = 0$ , for  $1 \leq i \leq \ell$ .

Now suppose that  $1 \leq i \leq \ell$  and that

$$f^{j_i} = z^{d \cdot j_i} - \zeta_{d \cdot j_i - 1} \lambda^{d \cdot j_i - 1} - \zeta_{d \cdot j_i - 2} \lambda^{d \cdot j_i - 2} - \dots - \zeta_1 \lambda - \zeta_0.$$

Since  $f$  is irreducible with degree  $d \geq 2$ ,  $z$  does not divide either  $f$  or  $f^{j_i}$ , so that  $\zeta_0 \neq 0$ .

Now

$$(C_{f^{j_i}} - \lambda I_{d \cdot j_i})x_i = \begin{bmatrix} -\lambda & & & & 0 & \zeta_0 \\ 1 & -\lambda & & & 0 & \zeta_1 \\ 0 & 1 & -\lambda & & 0 & \zeta_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -\lambda & \zeta_{d \cdot j_i - 2} \\ 0 & 0 & 0 & \dots & 1 & \zeta_{d \cdot j_i - 1} - \lambda \end{bmatrix} \cdot x_i = 0.$$

The top left submatrix of  $C_{f^{j_i}} - \lambda I_n$  with order  $d \cdot j_i - 1$  is nonsingular — it is lower triangular with the nonzero value  $-\lambda$  at each diagonal position. The bottom entry of  $x_i$  must therefore be nonzero. It now suffices to consider a vector

$$x_i = \begin{bmatrix} \lambda^{d \cdot j_i - 2} \alpha_0 \\ \lambda^{d \cdot j_i - 3} \alpha_1 \\ \vdots \\ \lambda \alpha_{d \cdot j_i - 3} \\ \alpha_{d \cdot j_i - 2} \\ \lambda^{d \cdot j_i - 1} \end{bmatrix}$$

for  $\alpha_0, \alpha_1, \dots, \alpha_{d \cdot j_i - 2} \in \mathbb{K}$ . Examining the bottom entries in the above vector, it is now possible to prove inductively that

$$\begin{aligned} \alpha_i &= \lambda^{d \cdot j_i} - \zeta_{d \cdot j_i - 1} \lambda^{d \cdot j_i - 1} - \zeta_{d \cdot j_i - 2} \lambda^{d \cdot j_i - 2} - \dots - \zeta_{i+2} \lambda^{i+2} - \zeta_{i+1} \lambda^{i+1} \\ &= \zeta^i \lambda^i + \zeta^{i-1} \lambda^{i-1} + \dots + \zeta_1 \lambda + \zeta_0. \end{aligned}$$



Consequently

$$x_i = \begin{bmatrix} \lambda^{d \cdot j_i - 2} \zeta_0 \\ \lambda^{d \cdot j_i - 2} \zeta_1 + \lambda^{d \cdot j_i - 3} \zeta_0 \\ \vdots \\ \lambda^{d \cdot j_i - 2} \zeta_{d \cdot j_i - 2} + \cdots + \lambda \zeta_1 + \zeta_0 \\ \lambda^{d \cdot j_i - 1} \end{bmatrix}$$

It follows (considering the powers of  $\lambda$  with coefficient  $\zeta_0$ , above) that if the entries of the vector  $u \in \mathbb{F}^{1 \times n}$  are chosen uniformly and independently from  $\mathbb{F}$  then  $u \cdot \hat{x}_i = g(\lambda)$  where  $g$  is a uniformly chosen polynomial in  $\mathbb{F}[z]$  with degree at most  $d \cdot j_i - 1$ . Consequently, since  $g(\lambda) = 0$  only if  $g$  is divisible by  $f$ ,  $(A - \lambda I_n + v \cdot u) \hat{x}_i = 0$  with probability at most  $|\mathbb{F}|^{-d}$ . Furthermore, the events that  $(A - \lambda I_n + v \cdot u) \hat{x}_1 = 0, (A - \lambda I_n + v \cdot u) \hat{x}_2 = 0, \dots, (A - \lambda I_n + v \cdot u) \hat{x}_\ell = 0$  are mutually independent, since these involve pairwise disjoint subsets of the entries of  $u$ .

It follows that the conditional probability that  $A + v \cdot u$  has exactly  $\ell - 1$  invariant factors divisible by  $f$ , given that  $v$  is not in the column space of  $A - \lambda I_n$ , is at least  $(1 - |\mathbb{F}|^{-\ell d})$ . Thus  $A + v \cdot u$  has exactly  $\ell - 1$  invariant factors divisible by  $f$  with probability at least  $(1 - |\mathbb{F}|^{-\ell d})^2$ , as needed to complete the proof of part (a) of the claim.

In order to prove part (b), suppose that the minimal polynomial of  $A$  is not divisible by  $f$ , so that the matrix  $A - \lambda I_n$  is a nonsingular matrix in  $\mathbb{K}^{n \times n}$ . In this case there exist matrices  $B_0, B_1, \dots, B_{d-1} \in \mathbb{F}^{n \times n}$  such that

$$(A - \lambda I_n)^{-1} = B_0 + \lambda B_1 + \lambda^2 B_2 + \cdots + \lambda^{d-1} B_{d-1}. \quad (5.8)$$

It follows that

$$\begin{aligned} I_n &= (A - \lambda I_n)(B_0 + \lambda B_1 + \lambda^2 B_2 + \cdots + \lambda^{d-1} B_{d-1}) \\ &= AB_0 + \gamma_0 B_{d-1} + \sum_{i=1}^{d-1} (AB_i - B_{i-1} + \gamma_i B_{d-1}) \lambda^i, \end{aligned}$$

so that  $AB_0 + \gamma_0 B_{d-1} = I_n$  and  $B_{i-1} = AB_i + \gamma_i B_{d-1}$  for  $1 \leq i \leq d-1$ . It can now be proved inductively, using the above equations, that

$$B_{d-i} = A^{i-1} B_{d-1} + \sum_{j=d-i+1}^{d-1} \gamma_j A^{j-d+i-1} B_{d-1} \quad (5.9)$$

for every integer  $i$  such that  $1 \leq i \leq d$ .

Now let  $v \in \mathbb{F}^{n \times 1}$ . Suppose that  $f$  divides the minimal polynomial of  $A + v \cdot u$ , so that the matrix  $A - \lambda I_n + v \cdot u$  is singular in  $\mathbb{K}^{n \times n}$ . There must exist a nonzero vector  $x \in \mathbb{K}^{n \times 1}$  such that  $(A - \lambda I_n + v \cdot u)x = 0$ . In this case  $(A - \lambda I_n)x = -v \cdot (u \cdot x)$ , so that  $(A - \lambda I_n)x$  is a  $\mathbb{K}$ -linear multiple of  $v$ . Since  $(A - \lambda I_n + v \cdot u)x = 0$  if and only if

$(A - \lambda I_n + v \cdot u)(\alpha x) = 0$  for any nonzero element  $\alpha$  of  $K$ , it now suffices to consider the vector  $x = -(A - \lambda I_n)^{-1}v$  — so that  $u(A - \lambda I_n)^{-1}v = 1$ .

Since  $(A - \lambda I_n)^{-1}$  is as shown at line (5.8), above, it now follows that  $u_i B_0 v = 1$  and  $u_i B_i v = 0$  for  $1 \leq i \leq d-1$ . It now follows by the equation at line (5.9) that  $uA^j B_{d-1} v = 0$  for  $0 \leq j \leq d-2$  and that  $uA^{d-1} B_{d-1} v = 1$ .

Consider the minimal polynomial of the matrix  $A$  and the vector  $B_{d-1}v$ , for  $B_{d-1}$  as above — that is, the monic polynomial  $g \in F[z]$  with least degree such that  $g(A)B_{d-1} = 0 \in F^{n \times 1}$ . If the degree of this polynomial is less than  $d$  then there is no vector  $u \in F^{1 \times n}$  such the above conditions are satisfied — for  $A^{d-1}B_{d-1}v$  is a linear combination of  $B_{d-1}v, AB_{d-1}v, \dots, A^{d-2}B_{d-1}v$  in this case, and  $uA^{d-1}B_{d-1}v = 0$  if  $A^i B_{d-1}v = 0$  for  $0 \leq i \leq d-2$ .

On the other hand, if the degree of this minimal polynomial is at least  $d$  then the vectors  $B_{d-1}v, AB_{d-1}v, \dots, A^{d-1}B_{d-1}v$  are linearly independent, so that the matrix  $C \in F^{n \times d}$  with these vectors as columns has maximal rank  $d$ . It now suffices to note that the vector  $u \in F^{1 \times n}$  only satisfies the condition required above if

$$uC = \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \end{bmatrix} \in F^{1 \times d}.$$

Since the entries of  $u$  are chosen uniformly and independently from those of  $v$  it now follows that the probability that  $(A - \lambda I_n + v \cdot u)x = 0$  — and that  $A - \lambda I_n + v \cdot u$  is singular — is at most  $|F|^{-d}$ , establishing part (b) of the claim.  $\square$

**Theorem 5.8.** *Let  $A \in F^{n \times n}$  and let  $f \in F[z]$  be a monic irreducible polynomial with degree  $d$  such that at most  $k$  invariant factors of  $A$  are divisible by  $f$  for some positive integer  $k$ . Suppose the entries of matrices  $U \in F^{k \times n}$  and  $V \in F^{n \times k}$  are chosen uniformly and independently from  $F$ .*

- (a) *If exactly  $k$  of the invariant factors of  $A$  are divisible by  $f$  then the probability that the minimal polynomial of  $A + V \cdot U$  is not divisible by  $f$  is at least*

$$\left(1 - \frac{1}{|F|^{2d} - 1}\right) (1 - |F|^{-d})^2 = \frac{(|F|^{2d} - 2)(|F|^d - 1)}{|F|^{2d}(|F|^d + 1)} \geq 1 - 2|F|^{-d}.$$

- (b) *If the minimal polynomial of  $A$  is not divisible by  $f$  then the probability that the minimal polynomial of  $A + V \cdot U$  is divisible by  $f$  is at least*

$$\frac{|F|^d - 1}{|F|^d} = 1 - |F|^{-d}$$

*if  $k = 1$ , and with probability at least*

$$\left(1 - \frac{|F|^{-d}}{|F|^d - 1}\right) (1 - |F|^{-d}) = 1 - \frac{|F|^d + 1}{|F|^{2d}} > 1 - 2|F|^{-d}$$

*if  $k \geq 2$ .*

(c) *The minimal polynomial of  $A + V \cdot U$  is not divisible by  $f$  with probability at least*

$$\frac{(|F|^{4d} - 2)(|F|^{2d} - |F|^d - 1)}{|F|^{3d}(|F|^{3d} + |F|^{2d} + |F|^d + 1)} = 1 - \frac{2|F|^{5d} + 2|F|^{4d} + |F|^{3d} + 2|F|^{2d} - 2|F|^d - 2}{|F|^{3d}(|F|^{3d} + |F|^{2d} + |F|^d + 1)} \geq 1 - 2|F|^{-d}.$$

*Proof.* The proof of this result is virtually identical to the proof of Lemma 5.4. Rather than considering the rank of a sequence of matrices, one should consider the number of invariant factors of a sequence of matrices that are divisible by the polynomial  $f$ . Lemma 5.7 replaces Lemma 5.3 in the argument, so that  $|F|$  is consistently replaced by  $|F|^d$  in the bounds that are being applied and derived.  $\square$

Theorem 5.2 now follows by Theorem 5.6 and part (c) of Theorem 5.8.

## 6 Nontrivial Nilpotent Blocks

Recall that a nilpotent block in the Jordan form of a matrix  $A \in \mathbb{F}^{n \times n}$  is **nontrivial** if has order at least two — so that the minimal polynomial of this block is  $z^j$  for  $j \geq 2$ . The purpose of this section is to establish the following.

**Theorem 6.1.** *It is possible to decide whether a matrix  $A \in \mathbb{F}^{n \times n}$  has at most  $k$  nontrivial nilpotent blocks, in such a way that the incorrect decision is reached with probability at most  $\epsilon$  for any positive constant  $\epsilon$ . The cost to do so includes the selection of  $\Theta(nk)$  values uniformly and independently from  $\mathbb{F}$  and  $\Theta(n^2k + \mu n)$  arithmetic operations in  $\mathbb{F}$ .*

*It is also possible both to certify that  $A$  has at most  $k$  nontrivial nilpotent blocks and to certify that  $A$  has more than  $k$  nontrivial nilpotent blocks. In both cases the verifier is guaranteed to accept if the prover's information is correct. The verifier accepts with probability at most  $\epsilon$  if the prover's information is incorrect.*

*For both protocols, the expected cost for the prover to complete the protocol is dominated by the worst-case cost for the initial decision stage, as given above. The cost for the verifier, when confirming that  $A$  has at most  $k$  nontrivial nilpotent blocks, includes the selection of  $\Theta(n)$  values, uniformly and independently from  $\mathbb{F}$ , and  $\Theta(nk + \mu)$  arithmetic operations in  $\mathbb{F}$ . The cost for the verifier, when confirming that  $A$  has more than  $k$  nontrivial nilpotent blocks, includes the selection of  $\Theta(nk)$  values, uniformly and independently from  $\mathbb{F}$ , and  $\Theta(nk + \mu)$  arithmetic operations in  $\mathbb{F}$ .*

### 6.1 Detection

Let  $\sigma_1(2) = 17$ ,  $\sigma_1(3) = 3$ ,  $\sigma_1(4) = \sigma_1(5) = 2$ ,  $\sigma_1(q) = 1$  when  $q = 7$  and  $\sigma_1(q) = 2$  when  $q \geq 8$ . It is easily checked that if  $\rho_1(q)$  is as given in Theorem 5.2, for every prime power  $q$ , then  $(1 - \rho_1(q))^{\sigma_1(q)} \leq 1/2$  when  $2 \leq q \leq 7$  and  $(1 - \rho_1(q))^{\sigma_1(q)} \leq q^{-1}$  when  $q \geq 8$ .

In order to check whether a black-box matrix  $A \in \mathbb{F}^{n \times n}$  has at most  $k$  nontrivial invariant factors, when  $\mathbb{F} = \mathbb{F}_q$ , and to ensure that the probability of an incorrect decision is at most a positive constant  $\epsilon$ , the prover should select  $\tau$  pairs of matrices  $U_i \in \mathbb{F}^{k \times n}$  and  $V_i \in \mathbb{F}^{n \times k}$ , for  $1 \leq i \leq \tau$ , by choosing the entries of these matrices uniformly and independently from  $\mathbb{F}$  — where  $\tau = \lceil \log_2(2 \cdot \epsilon^{-1}) \rceil \cdot \sigma_1(q)$  if  $2 \leq q \leq 7$  and  $\tau = \lceil \log_q(2 \cdot \epsilon^{-1}) \rceil \cdot \sigma_1(q)$  if  $q \geq 8$ .

Suppose first that  $A$  has more than  $k$  nontrivial nilpotent blocks, so that the  $k + 1^{\text{st}}$  invariant factor of  $A$  is divisible by  $z^2$ . Then it follows by Theorem 5.1, above, that the minimal polynomial of  $A + V_i \cdot U_i$  is divisible by  $z^2$  for all  $i$ .

For every such matrix it is easily checked, in this case, that if  $u_{i,j}, v_{i,j} \in \mathbb{F}^{n \times 1}$  are chosen uniformly and independently from  $\mathbb{F}^{n \times 1}$  then the minimal polynomial of the linearly recurrent sequence

$$u_{i,j}^T v_{i,j}, u_{i,j}^T (A + V_i \cdot U_i) v_{i,j}, u_{i,j}^T (A + V_i \cdot U_i)^2 v_{i,j}, \dots \quad (6.1)$$

is also divisible by  $z^2$  with probability at least  $(1 - 1/q)^2$  — which is equal to  $1/4$  if  $q = 2$ , and greater than  $1 - 2/q$  if  $q \geq 3$ . Consequently, if  $\lambda = \lceil \log_{4/3}(2 \cdot \tau \cdot \epsilon^{-1}) \rceil$  when  $q = 2$ , and  $\lambda = \lceil \log_{q/2}(2 \cdot \tau \cdot \epsilon^{-1}) \rceil$  when  $q \geq 3$ , and pairs of vectors  $u_{i,j}$  and  $v_{i,j}$  are chosen uniformly and independently from  $\mathbb{F}^{n \times 1}$ , for  $1 \leq i \leq \tau$  and  $1 \leq j \leq \lambda$ , then, for each  $i$ , the probability there is no integer  $j$  such that  $1 \leq j \leq \lambda$ , and the minimal polynomial of the linear recurrence at line (6.1) is divisible by  $z^2$ , is at most  $\epsilon/(2\tau)$ . The probability that it has not been confirmed that the minimal polynomial of  $A + V_i \cdot U_i$  is divisible by  $z^2$ , for all  $i$  such that  $1 \leq i \leq \tau$ , is therefore certainly at most  $\epsilon/2 < \epsilon$  in this case.

It follows that — for fixed  $q$  and  $\epsilon$  — the number of applications of Wiedemann's algorithm needed to compute the minimal polynomials of sequences as above and confirm the above condition, with the desired reliability, is a constant.

On the other hand, a straightforward calculation (involving  $\tau$ , as given above) and an application of Theorem 5.2(a) establishes that if  $A$  has at most  $k$  nontrivial invariant factors then the minimal polynomial of at least one matrix  $A + V_i \cdot U_i$ , such that  $1 \leq i \leq \tau$ , is *not* divisible by  $z^2$  with probability at least  $1 - \epsilon/2$ .

One should again try to compute the minimal polynomial of each matrix  $A + V_i \cdot U_i$  by computing the minimal polynomials of linearly recurrent sequences of the form shown at line (6.1) for  $\lambda$  uniformly and independently pairs of vectors  $u_{i,j}, v_{i,j} \in \mathbb{F}^{n \times 1}$ .

If  $1 \leq i \leq \tau$  and the minimal polynomial of  $A + V_i \cdot U_i$  is not divisible by  $z^2$ , then the minimal polynomial of the linear recurrent sequence shown at line (6.1) is not divisible by  $z^2$ , either, for any  $j$  that  $1 \leq j \leq \lambda$ .

On the other hand, it follows by the above analysis that there will exist an integer  $j$  such that  $1 \leq j \leq \lambda$  and the minimal polynomial of the above linearly recurrent sequence is divisible by  $z^2$ , for every integer  $i$  such that  $1 \leq i \leq \tau$  and the minimal polynomial of  $A + V_i \cdot U_i$  is divisible by  $z^2$ , with probability at least  $1 - \epsilon/2$ .

In this case a pair of matrices  $U \in \mathbb{F}^{k \times n}$  and  $V \in \mathbb{F}^{n \times k}$ , such that the minimal polynomial of  $A + V \cdot U$  is not divisible by  $z^2$ , can be selected by choosing any one of the pairs of

matrices  $U_i$  and  $V_i$  that have not been eliminated using the above process. The probability that either this case has not been correctly identified, or a pair of matrices  $U$  and  $V$  as described above has not been correctly selected, is at most  $\epsilon$ .

Since the cost to multiply  $A + V \cdot U$  by a vector  $v \in \mathbb{F}^{n \times 1}$ , includes the cost,  $\mu$ , to multiply  $A$  by a vector, along with  $\Theta(nk)$  additional operations, it is straightforward to modify the analysis of Wiedemann's algorithm in order to conclude that the cost of the above process includes the uniform and independent selection of  $\Theta(nk)$  values from  $\mathbb{F}$ , along with  $\Theta(n^2k + \mu n)$  arithmetic operations in  $\mathbb{F}$ , as claimed.

## 6.2 Few Nilpotent Blocks: Certification and Verification

If the prover has determined that  $A$  has at most  $k$  nontrivial nilpotent blocks, as described above, then the prover should **commit** by sending matrices  $U \in \mathbb{F}^{k \times n}$  and  $V \in \mathbb{F}^{n \times k}$ , such that the minimal polynomial of  $A + V \cdot U$  is not divisible by  $z^2$ , to the verifier: These have now been obtained.

It is easily checked, by consideration of a rational Jordan form, that if the minimal polynomial of a matrix  $B \in \mathbb{F}^{n \times n}$  is not divisible by  $z^2$ , then a system  $Bx = b$  is consistent (for a given vector  $b \in \mathbb{F}^{n \times 1}$ ), if and only if the system  $B^2x = b$  is consistent as well. On the other hand, if the minimal polynomial of  $B$  is divisible by  $z^2$  and a vector  $c \in \mathbb{F}^{n \times n}$  is selected uniformly and independently, then the probability that the system  $B^2x = Bc$  is consistent is at most  $|\mathbb{F}|^{-1}$ .

The verifier may therefore form a **challenge** by selecting  $\gamma = \lceil \log_q \epsilon^{-1} \rceil$  vectors  $c_1, c_2, \dots, c_\gamma$  uniformly and independently from  $\mathbb{F}^{n \times 1}$  (for  $q = |\mathbb{F}|$ ), computing  $b_i = (A + V \cdot U)c_i$  for  $1 \leq i \leq \gamma$ , and sending  $b_1, b_2, \dots, b_\gamma$  to the prover.

The prover should compute vectors  $x_1, x_2, \dots, x_\gamma \in \mathbb{F}^{n \times 1}$  such that  $(A + V \cdot U)^2 x_i = b_i$  (possibly by applying Wiedemann's algorithm twice, for each  $i$ ) and send these to the verifier. Finally, the verifier should check whether the required equalities are satisfied — **accepting** if they are, and **rejecting** otherwise.

If the prover's information is correct then the verifier accepts with certainty; otherwise the verifier accepts, incorrectly, with probability at most  $\epsilon$ . The cost to the prover, to complete this protocol, is dominated by the cost of the “detection” stage described above. Since the verifier must only choose  $\Theta(n)$  values uniformly and independently from  $\mathbb{F}$  and multiply a constant number of vectors by  $A + V \cdot U$ , the number of operations used by the verifier is as claimed.

## 6.3 Many Nilpotent Blocks: Certification and Verification

If the prover has determined, instead, that  $A$  has more than  $k$  nontrivial nilpotent blocks, then the prover should **commit** by advising the verifier of this.

The verifier should then select matrices  $U_i \in \mathbb{F}^{k \times n}$  and  $V_i \in \mathbb{F}^{n \times k}$  for  $1 \leq i \leq \tau$ , for  $\tau$  as above, by selecting the entries of these matrices uniformly and independently from  $\mathbb{F}$ .

These matrices should then be sent to the prover as a **challenge**.

In **response** the prover should return vectors  $x_i \in \mathbb{F}^{n \times 1}$  such that  $(A + V_i \cdot U_i)x_i \neq 0 = (A + V_i \cdot U_i)^2 x_i = 0$  for  $1 \leq i \leq \tau$ . The verifier should then **accept** if these conditions are all satisfied and **reject** otherwise.

Once again, Theorem 5.1 can be used to establish that this protocol is perfectly complete — the verifier accepts with certainty if the prover’s information is correct. Theorem 5.2 and a straightforward calculation establishes that it is also sound: If the prover’s information is incorrect then the probability that the verifier accepts is at most  $\epsilon$ .

In order to see that the additional cost for the prover is low, recall that the **minimal polynomial** of a matrix  $B \in \mathbb{F}^{n \times n}$  and vector  $v \in \mathbb{F}^{n \times 1}$  is the monic polynomial  $g \in \mathbb{F}[z]$  with least degree such that  $g(B)v = 0$ . Wiedemann’s algorithm can be used to compute the minimal polynomial of  $A + V_i \cdot U_i$  and a vector  $v$ , as the least common multiple of a number of linear recurrent sequences as shown at line (6.1), with  $v = v_{i,j}$  and uniform and independent choices of the vector  $u_{i,j}$ . A small number of choices of  $u_{i,j}$  suffice to ensure that the minimal polynomial of  $(A + V_i \cdot U_i)v$  has been discovered with high probability. Furthermore, if vectors  $v_{i,j} \in \mathbb{F}^{n \times 1}$  are chosen uniformly and independently from  $\mathbb{F}^{n \times 1}$ , for  $j = 1, 2, \dots$ , and the minimal polynomial of  $A + V_i \cdot U_i$  is divisible by  $z^2$ , then the expected number of vectors  $v_{i,j}$  that must be considered, before a vector  $v = v_{i,j}$  is found such that the minimal polynomial of  $A + V_i \cdot U_i$  and  $v$  is also divisible by  $z^2$ , is at most two.

Suppose now that a vector  $v \in \mathbb{F}^{n \times 1}$  has been discovered and it has been confirmed that the minimal polynomial of  $A + V_i \cdot U_i$  and  $v$  is  $z^2 g$  for some polynomial  $g \in \mathbb{F}[z]$ . It suffices to compute and return the vector  $x_i = g(A + V_i)v$  in order to satisfy the requirements given above.

The expected cost for the prover to complete this protocol is, once again, dominated by the worst-case cost of the “detection” stage. The cost for the verifier includes the selection of  $\Theta(nk)$  values uniformly and independently from  $\mathbb{F}$  and a constant number of multiplications of  $A + V_i \cdot U_i$  by vectors, for matrices  $U_i \in \mathbb{F}^{k \times n}$  and  $V_i \in \mathbb{F}^{n \times k}$  — establishing the above claim.

## 7 Nontrivial Invariant Factors

The purpose of this section is to establish the following.

**Theorem 7.1.** *One can decide whether a matrix  $A \in \mathbb{F}^{n \times n}$  has at most  $k$  nontrivial invariant factors, such that the incorrect decision is made with probability at most  $\epsilon$  for any positive constant  $\epsilon$ . The expected cost of this includes the selection of  $\Theta(nk)$  values uniformly and independently from  $\mathbb{F}$  and  $\Theta(n^2k + \mu n)$  arithmetic operations in  $\mathbb{F}$ .*

*It is also possible both to certify that  $A$  has at most  $k$  nontrivial invariant factors and to certify that  $A$  has more than  $k$  nontrivial invariant factors. In both cases the verifier is guaranteed to accept if the prover’s information is correct. The verifier accepts with probability at most  $\epsilon$  if the prover’s information is incorrect.*

For both protocols, the expected cost for the prover to complete the protocol is in

$$O(n^2 \mathcal{M}(n) + n^2 k \log_2 n + \mu n \log_2 n),$$

where  $\mathcal{M}(n)$  is the number of operations in  $\mathbb{F}$  required for an arithmetic operation in a extension  $\mathbb{E}$  with degree in  $O(\log_2 n)$  over  $\mathbb{F}$ . When certifying that  $A$  has at most  $k$  nontrivial invariant factors, the verifier selects  $O(n \log_2 n)$  values uniformly and independently from  $\mathbb{F}$  and performs  $\Theta(n \mathcal{M}(n) + nk + \mu)$  arithmetic operations in  $\mathbb{F}$ . When certifying that  $A$  has more than  $k$  nontrivial invariant factors the verifier selects  $O(n \log_2 n + nk)$  values uniformly and independently from  $\mathbb{F}$  and performs  $O(n \mathcal{M}(n) + nk + \mu)$  arithmetic operations in  $\mathbb{F}$ .

## 7.1 Detection

The  $k + 1^{\text{st}}$  invariant factor  $\varphi_{k+1}$  of  $A$  is divisible by  $z^2$  if and only if  $A$  has more than  $k$  nontrivial nilpotent blocks. The process described in Subsection 6.1 should be applied to check this, with parameters chosen to ensure that the probability of failure is at most  $\epsilon/2$ .

If  $\varphi_{k+1}$  is not divisible by  $z^2$  then a pair of matrices  $U_0 \in \mathbb{F}^{k \times n}$  and  $V_0 \in \mathbb{F}^{n \times k}$  have been found such that the minimal polynomial  $f_0 \in \mathbb{F}[z]$  of  $A + V_0 \cdot U_0$  is not divisible by  $z^2$  — and  $f_0$  has been correctly computed — with probability at least  $1 - \epsilon/2$ . The detection stage should proceed with an attempt to compute a factor  $\chi \neq z$  of  $\varphi_{k+1}$  with positive degree, along with a certificate of this factor — or to determine that no such factor exists.

Suppose first that  $q = |\mathbb{F}| = 2$ . In this case a straightforward variant of the protocol described in Subsection 6.1 can be used either to conclude that  $\varphi_{k+1}$  is divisible by  $z + 1$  or to obtain matrices  $U_1 \in \mathbb{F}^{k \times n}$  and  $V_1 \in \mathbb{F}^{n \times k}$  such that the minimal polynomial  $f_1 \in \mathbb{F}[z]$  of  $A + V_1 \cdot U_1$  is not divisible by  $z + 1$ . Suppose that this is carried out in such a way that the probability of failure is at most  $\epsilon/6$  — so that  $U_1$ ,  $V_1$  and  $f_1$  have also been obtained with probability at least  $1 - \epsilon/6$  as well in the second case. If  $\sigma_2(2, 1) = 6$  and  $\rho_2(q, d)$  is as shown in part (b) of Theorem 5.2 then  $(1 - \rho_2(2, 1))^{\sigma_2(2, 1)} < \frac{1}{2}$ . It therefore suffices to choose  $\tau_2(2, 1) = \lceil \log_2(12/\epsilon) \rceil \cdot \sigma_2(2, 1)$  pairs of matrices  $\hat{U}_i \in \mathbb{F}^{k \times n}$  and  $\hat{V}_i \in \mathbb{F}^{n \times k}$ , for  $1 \leq i \leq \tau_2(2, 1)$ , in order to ensure that the minimal polynomial of  $A + \hat{V}_i \cdot \hat{U}_i$  is not divisible by  $z + 1$ , for at least one of these pairs of matrices  $\hat{U}_i$  and  $\hat{V}_i$ , with probability at least  $1 - \epsilon/12$ , if  $\varphi_{k+1}$  is not divisible by  $z + 1$ . If  $\lambda_2(2, 1) = \lceil \log_{4/3}(12 \cdot \tau_2(2, 1) \cdot \epsilon^{-1}) \rceil$  then a consideration of the minimal polynomials of  $\lambda_2(2, 1)$  linearly recurrent sequences with the form shown at line (6.1), for  $1 \leq i \leq \tau_2(2, 1)$ , suffice to ensure that it has been correctly discovered whether  $z + 1$  divides the minimal polynomial of  $A + \hat{V}_i \cdot \hat{U}_i$ , for all of these matrices, with probability at least  $1 - \epsilon/12$  — as is necessary and sufficient here.

If it has been determined that  $\varphi_{k+1}$  is not divisible by  $z + 1$ , then a variant of the above protocol should be applied, once again, either to conclude that  $\varphi_{k+1}$  is divisible by  $z^2 + z + 1$  (the only monic irreducible polynomial in  $\mathbb{F}[z]$  with degree two) or to obtain matrices  $U_2 \in \mathbb{F}^{k \times n}$  and  $V_2 \in \mathbb{F}^{n \times k}$  such that the minimal polynomial  $f_2 \in \mathbb{F}[x]$  of  $A + V_2 \cdot U_2$  is not divisible by  $z^2 + z + 1$ . Suppose, as above, that this is carried out in such a way

that the probability of failure (including failure to compute  $U_2$ ,  $V_2$  and  $f_2$ ) is at most  $\epsilon/6$ . Now, if  $\rho_2(q, d)$  is as given in part (b) of Theorem 5.2 then  $1 - \rho_2(2, 2) < \frac{1}{2}$ . It therefore suffices to choose  $\tau_2(2, 2) = \lceil \log_2(12/\epsilon) \rceil$  pairs of matrices  $\widehat{U}_i \in \mathbb{F}^{k \times n}$  and  $\widehat{V}_i \in \mathbb{F}^{n \times k}$ , for  $1 \leq i \leq \tau_2(2, 2)$ , in order to ensure that the minimal polynomial of  $A + \widehat{V}_i \cdot \widehat{U}_i$  is not divisible by  $z^2 + z + 1$ , for at least one of these pairs of matrices  $\widehat{U}_i$  and  $\widehat{V}_i$ , with probability at least  $1 - \epsilon/12$  if  $\varphi_{k+1}$  is not divisible by  $z^2 + z + 1$ .

Suppose now that  $\widehat{U}_i \in \mathbb{F}^{k \times n}$  and  $\widehat{V}_i \in \mathbb{F}^{n \times k}$  such that the minimal polynomial of  $A + V_i \cdot U_i$  is divisible by  $z^2 + z + 1$ . Then, if vectors  $u_{i,j}, v_{i,j}$  are chosen uniformly and independently from  $\mathbb{F}^{n \times 1}$ , then the minimal polynomial of the linear recurrence resembling that shown at line (6.1) (with  $\widehat{U}_i$  and  $\widehat{V}_i$  replacing  $U_i$  and  $V_i$ , respectively) is divisible by  $z^2 + z + 1$  with probability at least  $(1 - \frac{1}{4})^2 = \frac{9}{16}$ . Consequently, if  $\lambda_2(2, 2) = \lceil \log_{16/7}(12 \cdot \tau_2(2, 2) \cdot \epsilon^{-1}) \rceil$ , then a consideration of  $\lambda_2(2, 2)$  linearly recurrent sequences resembling the one at line (6.1), for  $1 \leq i \leq \tau_2(2, 2)$ , suffices to ensure that it has been correctly discovered whether  $z^2 + z + 1$  divides the minimal polynomial of  $A + \widehat{V} \cdot \widehat{U}$ , for all of these matrices, with probability at least  $1 - \epsilon/12$  — as is necessary and sufficient, here, once again.

Suppose, now, that it has been determined that  $\varphi_{k+1}$  is not divisible by  $z^2 + z + 1$  either. Suppose that an additional two pairs of matrices  $U_i \in \mathbb{F}^{k \times n}$  and  $V_i \in \mathbb{F}^{n \times k}$  are selected uniformly and independently, for  $3 \leq i \leq 4$ . Then, since  $1 - \rho_2(2, d) \leq 2^{1-d}$  for  $d \geq 3$ , and there are only two monic irreducible polynomials in  $\mathbb{F}[z]$  with degree three, while there are at most  $q^d/d = 2^d/d$  monic irreducible polynomials with degree  $d$  in  $\mathbb{F}[z]$  when  $d \geq 4$ , it follows that if  $f_i$  is the minimal polynomial of  $A + V_i \cdot U_i$ , for  $3 \leq i \leq 4$ , then the probability that  $\gcd(f_3, f_4)$  has a monic irreducible factor, with degree at least three, that is not also a factor of  $\varphi_{k+1}$ , is at most

$$\begin{aligned}
& 2 \times (1 - \rho_2(2, 3))^2 + \sum_{d \geq 4} \left( \frac{2^d}{d} \right) \cdot (1 - \rho_2(2, d))^2 \\
& \leq 2 \times \left( \frac{1}{4} \right)^2 + \sum_{d \geq 4} \left( \frac{2^d}{d} \right) \cdot \left( \frac{2}{2^d} \right)^2 \quad (\text{by the bounds in part (b) of Theorem 5.2}) \\
& \leq \frac{1}{8} + \sum_{d \geq 4} \left( \frac{2^d}{4} \right) \cdot \left( \frac{2}{2^d} \right)^2 \\
& = \frac{1}{8} + \sum_{d \geq 4} 2^{-d} \\
& = \frac{1}{8} + \frac{1}{8} \\
& = \frac{1}{4}.
\end{aligned}$$

Consequently if one chooses matrices  $U_i \in \mathbb{F}^{k \times n}$  and  $V_i \in \mathbb{F}^{n \times k}$  uniformly and independently, for  $3 \leq i \leq \tau + 2$ , instead, where  $\tau = 2 \lceil \log_4(12\epsilon^{-1}) \rceil$ , and  $f_i$  is the minimal



polynomial of  $A + V_i \cdot U_i$  for all such  $i$ , then  $\gcd(f_3, f_4, \dots, f_{\tau+2})$  has an irreducible factor, with degree at most three, that is not also a factor of  $\varphi_{k+1}$ , with probability at most  $\epsilon/12$ . Indeed, two pairs of matrices  $U_i \in \mathbb{F}^{k \times n}$  and  $V_i \in \mathbb{F}^{n \times k}$  will have been identified, with probability at least  $1 - \epsilon/12$ , such that if  $f_i$  is the minimal polynomial of  $U + V_i \cdot U_i$ , for  $3 \leq i \leq 4$  and  $f_0, f_1$  and  $f_2$  are as above, then the squarefree part of  $\gcd(f_0, f_1, f_2, f_3, f_4)$  is a divisor of  $\varphi_{k+1}$ .

As discussed in Subsection 6.1, and above, it is possible to ensure that each of the above minimal polynomials  $f_i$  of  $A = V_i \cdot U_i$  is computed in such a way that the probability of failure here is also at most  $\epsilon/12$ , by computing the minimal polynomials of  $\Theta(\log_2(\epsilon^{-1}))$  linearly recurrent sequences as shown at line (6.1). At this point, either a divisor of  $\varphi_{k+1}$  with positive degree that is different from  $z$  has been identified, or matrices  $U_i \in \mathbb{F}^{k \times n}$ ,  $V_i \in \mathbb{F}^{n \times k}$ , and the minimal polynomials  $f_i \in \mathbb{F}[z]$  of  $A + V_i \cdot U_i$  have been identified, for  $0 \leq i \leq 4$ , such that  $\gcd(f_0, f_1, f_2, f_3, f_4) = \varphi_{k+1} \in \{1, z\}$ . The probability of failure of this process is at most  $\epsilon$ , and (for fixed  $\epsilon$ ) the prover has selected  $\Theta(nk)$  values uniformly and independently from  $\mathbb{F}$ , and performed  $\Theta(n^2k + \mu n)$  arithmetic operations in  $\mathbb{F}$ .

Suppose next that  $q \geq 3$ . In this case one should begin, once again, by applying the process described in Section 6 to determine whether  $A$  has more than  $k$  nontrivial nilpotent blocks, in such a way that this process fails with probability at most  $\epsilon/2$  — and in such a way that a pair of matrices  $U_0 \in \mathbb{F}^{k \times n}$  and  $V_0 \in \mathbb{F}^{n \times k}$  has been discovered such that the minimal polynomial  $f_0$  of  $A + V_0 \cdot U_0$  is not divisible by  $z^2$  — and  $f_0$  has been computed — if the process has not failed and  $A$  has at most  $k$  nontrivial nilpotent blocks.

Suppose now that  $A$  has at most  $k$  nontrivial nilpotent blocks, so that it is necessary to check whether  $\varphi_{k+1}$  has a monic irreducible factor in  $\mathbb{F}[z]$  that is different from  $z$ . Let  $c$  be a positive integer, greater than or equal to two, and suppose that  $c$  pairs of matrices  $U_i \in \mathbb{F}^{k \times k}n$  and  $V_i \in \mathbb{F}^{n \times n}k$  are chosen uniformly and independently, for  $1 \leq i \leq c$ .

Since there are  $q - 1$  monic irreducible polynomials with degree in  $\mathbb{F}[z]$  with degree one, it follows by part (b) of Theorem 5.2 that  $\gcd(f_1, f_2, \dots, f_c)$  has a monic irreducible factor that is not also a factor of  $\varphi_{k+1}$  with probability at most

$$g_1(q, c) = (q - 1) \cdot (1 - \rho_2(q, 1))^c.$$

Since there are at most  $\frac{q^d}{d}$  monic irreducible polynomials with degree  $d$  in  $\mathbb{F}[x]$ , for  $d \geq 2$ , it also follows that  $\gcd(f_1, f_2, \dots, f_c)$  has a monic quadratic irreducible factor in  $\mathbb{F}[z]$  that is not also a factor of  $\varphi_{k+1}$  with probability at most

$$g_2(q, c) = \frac{q^2}{2} \cdot (1 - \rho_2(q, 2))^c,$$

and a monic cubic irreducible factor in  $\mathbb{F}[z]$  that is not a factor of  $\varphi_{k+1}$  with probability at most

$$g_3(q, c) = \frac{q^3}{3} \cdot (1 - \rho_2(q, 3))^c.$$

Finally, the probability that  $\gcd(f_1, f_2, \dots, f_c)$  has a monic irreducible factor in  $\mathbb{F}[z]$  with degree at least four, that is not also a factor of  $\varphi_{k+1}$ , is at most

$$\begin{aligned}
\sum_{d \geq 4} \frac{q^d}{d} \cdot (1 - \rho_2(q, d))^c &\leq \sum_{d \geq 4} \frac{q^d}{4} \cdot (1 - \rho_2(q, d))^c \\
&\leq \sum_{d \geq 4} \frac{q^d}{4} \cdot \left(\frac{2}{q^d}\right)^c \\
&= 2^{c-2} \sum_{d \geq 4} (q^{1-c})^d \\
&= 2^{c-2} \frac{q^{4(1-c)}}{1 - q^{1-c}} \\
&= 2^{c-2} \frac{q^{3(1-c)}}{q^{c-1} - 1} = g_4(q, c).
\end{aligned}$$

Consequently  $\gcd(f_1, f_2, \dots, f_c)$  has a monic irreducible factor in  $\mathbb{F}[z]$ , different from  $z$ , that is not also a factor of  $\varphi_{k+1}$ , is at most

$$F(q, c) = g_1(q, c) + g_2(q, c) + g_3(q, c) + g_4(q, c)$$

for the functions  $g_1, g_2, g_3$  and  $g_4$  that are given above.

Straightforward calculations, aided by the use of a computer algebra system, confirm that  $F(3, 4) < \frac{1}{2} < F(3, 3)$ , so that one can set  $c = 4$  when  $q = 3$  in order to ensure that the above process succeeds with probability at least  $\frac{1}{2}$  when  $q = 3$ . Similarly,  $F(q, 3) < \frac{1}{2} < F(q, 2)$ , so that one can set  $c = 3$ , when  $4 \leq q \leq 7$ . Finally,  $F(q, 2) < \frac{1}{2}$ , so that one can set  $c = 2$ , when  $q \geq 8$ .

While it seems necessary to set  $c = 2$  for large field sizes as well, the probability of failure drops (for  $c = 2$ ) as  $q = |\mathbb{F}|$  increases. Indeed, using the fact that  $(1 - \rho_2(q, d)) \leq 2q^{-d}$  for  $d \geq 1$ , it is straightforward to establish that  $F(q, 2) \leq 4q^{-1}$  when  $q \geq 5$ .

As for the case that  $q = 2$ ,  $\Theta(\log_2(\epsilon^{-1}))$  independent trials should be used, in order to ensure that a set of  $c$  pairs of matrices and corresponding minimal polynomials  $f_1, f_2, \dots, f_c$  have been found, so that the squarefree part of  $\gcd(f_0, f_1, \dots, f_c)$  is a divisor of  $\varphi_{k+1}$  with probability at least  $1 - \epsilon/4$ . Sufficiently many linear recurrences of the form shown at line (6.1) should be considered to ensure that all minimal polynomials of matrices  $A + V_i \cdot U_i$  have been correctly computed, with probability at least  $1 - \epsilon/4$  as well, in order to ensure that this “detection” process fails with probability at most  $\epsilon$ .

If it has been determined that  $A$  has at most  $k$  nontrivial invariant factors, so that  $c + 1$  pairs of matrices  $U_i \in \mathbb{F}^{k \times n}$  and  $V_i \in \mathbb{F}^{n \times k}$ , and corresponding minimal polynomials  $f_i \in \mathbb{F}[z]$  have been accumulated, such that  $\gcd(f_0, f_1, \dots, f_c) = \varphi_{k+1} \in \{1, z\}$ , then it will also be useful to compute polynomials  $g_0, g_1, \dots, g_c \in \mathbb{F}[z]$  such that

$$g_0 f_0 + g_1 f_1 + \dots + g_c f_c = \varphi_{k+1}. \quad (7.1)$$

Since  $c$  is a constant, for any field size  $q$ , the extended Euclidean algorithm can be applied to compute these polynomials, such that each polynomial  $g_i$  has degree in  $O(n)$ , at a cost that is dominated by the cost of the rest of this process.

## 7.2 Few Invariant Factors: Certification and Verification

In order to **commit** that  $A \in \mathbb{F}^{n \times n}$  has at most  $k$  invariant factors, the prover should send a sequence of  $c + 1$  pairs of matrices  $U_i \in \mathbb{F}^{k \times n}$  and  $V_i \in \mathbb{F}^{n \times k}$ , for  $0 \leq i \leq c$ , along with

- the minimal polynomial  $f_i \in \mathbb{F}[z]$  of  $A + V_i \cdot U_i$ ,
- polynomials  $g_i \in \mathbb{F}[z]$ , for  $0 \leq i \leq c$ , each with degree in  $O(n)$ , such that the equation at line (7.1) is also satisfied, where  $\varphi_{k+1} \in \{1, z\}$ .

As a **challenge**, the verifier should select  $\tau_3 = \lceil \log_{(2q-1-q^{-2})-1}(2(c+1)\tau_3\epsilon^{-1}) \rceil$  pairs of vectors  $u_i, v_i$  uniformly and independently from  $\mathbb{F}^{n \times 1}$ , for  $1 \leq i \leq \tau_3$ , and send these to the prover. The prover should then compute the minimal polynomial of each linearly recurrent sequence

$$u_i^T \cdot v_i, u_i^T (A + V_j \cdot U_j) \cdot v_i, u_i^T (A + V_j \cdot U_j)^2 v_i, \dots \quad (7.2)$$

for  $1 \leq i \leq \tau_3$  and  $0 \leq j \leq c$ , and should interact with the verifier to certify each of these minimal polynomials — choosing parameters in order to ensure that the verifier would discover an incorrect minimal polynomial with probability at least  $1 - \epsilon/(2(c+1)\tau_3)$  if the prover tried to provide one.

If the verifier can confirm that the minimal polynomial of each of the sequences shown at line (7.2) is divisible by the minimal polynomial  $f_i$  supplied by the prover, for  $0 \leq i \leq c$  and  $1 \leq j \leq \tau_3$  and, furthermore, can confirm that

$$g_0 f_0 + g_1 f_1 + \dots g_c f_c \in \{z, 1\},$$

then the verifier should **accept**. Otherwise, the verifier should **reject**.

Dumas, Kaltofen, Thomé and Villard [2] are primarily concerned with certification of the minimal polynomial of a matrix over a large field, namely a field  $\mathbb{F}$  such that  $|\mathbb{F}| \in \Omega(n)$ . However their observations about small field computations can be extended, in a straightforward way, to establish the following.

**Theorem 7.2** (Dumas, et. al.). *It is possible for a prover to compute the minimal polynomial of a given matrix  $A \in \mathbb{F}^{n \times n}$  and certify it using an interactive protocol that is perfectly complete and sound: An incorrect minimal polynomial is accepted with probability at most  $\epsilon$  for any desired positive constant  $\epsilon$ . The prover selects  $\Theta(n \log_2 n)$  values uniformly and independently from  $\mathbb{F}$  and performs  $\Theta(n^2 \mathcal{M}(n) + \mu n \log_2 n)$  additional operations in  $\mathbb{F}$  while participating in this process, where  $\mathcal{M}(n)$  is the number of operations in  $\mathbb{F}$  required*

for an arithmetic operation in a field extension whose degree over  $\mathbb{F}$  is logarithmic in  $n$ . The verifier selects  $O(n \log_2 n)$  values uniformly and independently from  $\mathbb{F}$  and performs  $O(n\mathcal{M}(n) + \mu)$  additional operations in  $\mathbb{F}$ .

Soundness and perfect completeness of this protocol are easily established, provided that the protocol of Dumas, Kaltofen, Thomé and Villard is used to ensure that any incorrect minimal polynomial of a linearly recurrent sequence as shown at line (7.2) would be detected by the verifier with probability at least  $1 - \frac{\epsilon}{2(c+1)\tau_3}$ . The costs for the prover and verifier, given in Theorem 7.1, follow from the fact that the cost to multiply  $A + V_i \cdot U_i$  by a vector is in  $\Theta(\mu + nk)$ .

### 7.3 Many Invariant Factors: Certification and Verification

In this case, the prover has detected a factor  $\chi \in \mathbb{F}[z]$  of  $\varphi_{k+1}$  that is different from 1 or  $z$ ; the prover should now **commit** to this protocol by sending  $\chi$  to the verifier. If  $\chi$  is divisible by  $z^2$  then verifier should interact with the prover in order to confirm that  $A$  has more than  $k$  nontrivial nilpotent blocks, as described in Section 6.

Otherwise, the verifier should choose matrices  $U_i \in \mathbb{F}^{k \times n}$  and  $V_i \in \mathbb{F}^{n \times k}$  uniformly and independently, for  $0 \leq i \leq \tilde{\tau} - 1$ , where  $\tilde{\tau} = \lceil \log_{(1-\rho_2(q,1))^{-1}}(2\epsilon^{-1}) \rceil$ , and send these to the prover as a **challenge**: If  $\varphi_{k+1}$  is not divisible by  $\chi$  then there will exist at least one pair of matrices  $U \in \mathbb{F}^{k \times n}$  and  $V \in \mathbb{F}^{n \times k}$ , such that the minimal polynomial of  $A + V \cdot U$  is also not divisible by  $\chi$ , with probability at least  $1 - \frac{\epsilon}{2}$ .

Consequently, as a **response** the prover should return vectors  $u_i, v_i \in \mathbb{F}^{n \times 1}$  such that the minimal polynomial of the linear recurrence

$$u_i^T v_i, u_i^T (A + V_i \cdot U_i) v_i, u_i^T (A + V_i \cdot U_i)^2 v_i, \dots$$

is  $\chi$ , for  $0 \leq i \leq \tilde{\tau} - 1$ . These can be obtained at sufficiently by following a process described, for example, by Eberly [3], to compute vectors  $u_i$  and  $v_i$  such the minimal polynomial of the above linearly recurrent sequence is the minimal polynomial  $f_i$  of  $A + V_i \cdot U_i$ , and then replacing  $v_i$  with  $(f_i/\chi)(A + V_i \cdot U_i)v_i$ .

The prover and verifier should apply the protocol of Dumas, Kaltofen, Thomé and Villard once again, in order to confirm this — ensuring that any incorrect pair of vectors is accepted with probability at most  $\epsilon/(2\tilde{\tau})$ , so that the total probability of failure is at most  $\epsilon$ , once again.

## 8 Additional Problems

Protocols certifying that a matrix  $A \in \mathbb{F}^{n \times n}$  is *not* banded, with band width  $k$ , or that the rank (or displacement rank, for the types of this discussed in Section 4) of a given matrix exceeds  $k$ , are also easily described as straightforward variants of those given in this report.

Additional properties allowing other “superfast” algorithms to be used might also be efficiently detected. For example, one might be able to detect some cases when *nested dissection* can be applied. The detection of *Vandermonde-like* and *Cauchy-like* matrices might be of interest — and also might be more challenging than that of detecting Toeplitz-like matrices: The operator matrices for Vandermonde-like and Cauchy-like matrices are defined using diagonal matrices whose entries can vary, and one would need to discover these diagonal matrices as part of a detection process.

This report has focussed on the case where  $k$  is extremely small. However, various “superfast” algorithms would still be superior to a black box algorithm for larger  $k$  — for example, for  $k \leq \sqrt{n}$ . Detection and conversion protocols that are effective, for larger  $k$ , might therefore be of interest.

Protocols such that the parameter  $k$  is not selected by a client (or verifier), but is instead discovered by the service provider (or prover), can be generally be obtained by modifying the protocols in this report in a straightforward way.

Finally, black box algorithms are also used for various integer matrix computations, including computations involving the *Smith form* of a matrices. Protocols to decide and certify whether the number of nontrivial elementary divisors of a given integer matrix would therefore be of interest.

## References

- [1] J.-G. Dumas and E. Kaltofen. Essentially optimal interactive certificates in linear algebra. In *Proceedings, 2014 International Symposium on Symbolic and Algebraic Computation (ISSAC '14)*, pages 146–153. ACM Press, 2014.
- [2] J.-G. Dumas, E. Kaltofen, E. Thomé, and G. Villard. Linear time interactive certificates for the minimal polynomial and determinant of a sparse matrix. In *Proceedings, 2016 International Symposium on Symbolic and Algebraic Computation (ISSAC '16)*, pages 199–206. ACM Press, 2016.
- [3] W. Eberly. Black box Frobenius decompositions over small fields. In *Proceedings, 2000 International Symposium on Symbolic and Algebraic Computation (ISSAC 2000)*, pages 106–113. ACM Press, 2000.
- [4] W. Eberly. Selecting algorithms for black box matrices: Checking for matrix properties that can simplify computations. In *Proceedings, 2016 International Symposium on Symbolic and Algebraic Computation (ISSAC '16)*, pages 207–214. ACM Press, 2016.
- [5] R. Frievālds. Fast probabilistic algorithms. In *Mathematical Foundations of Computer Science 1979*, volume 74 of *Lecture Notes in Computer Science*, pages 57–69. Springer-Verlag, 1979.

- [6] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, fourth edition, 2012.
- [7] V. Y. Pan. *Structured Matrices and Polynomials: Unified Superfast Algorithms*. Birkhäuser, 2001.
- [8] G. Villard. Computing the Frobenius normal form of a sparse matrix. In *Computer Algebra in Scientific Computing (CACS 2000)*, pages 395–407. Springer, 2000.